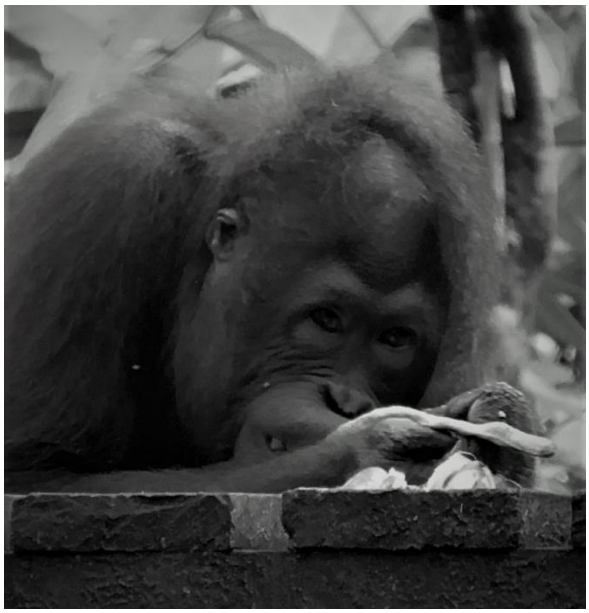
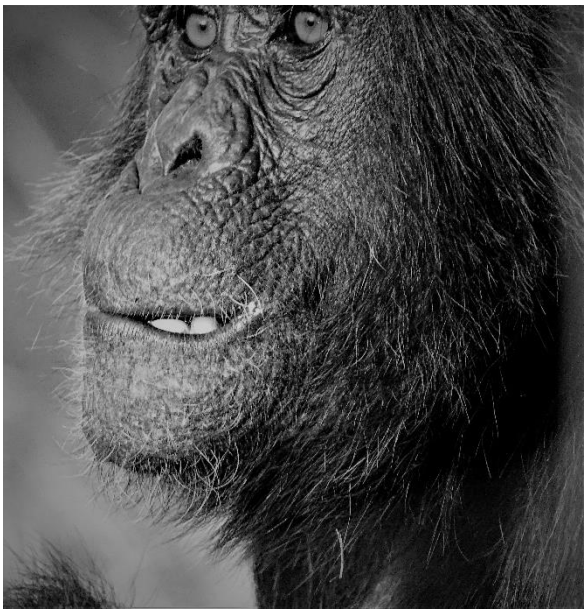
Measuring Personality in Great Apes: Special Focus on Naturally Occurring Contexts





Hélène Chotard

Thesis submitted in partial fulfilment of the requirements for the award of the degree of Doctor of Philosophy of the University of Portsmouth.

October 2019

University of Portsmouth

Abstract

Animals, like humans, differ as individuals from one another in the way they consistently express their behaviours across multiple contexts and time. In this thesis, I investigate the personality of sanctuary chimpanzees and rehabilitant orangutans. While leaning on an in-depth review of methodologies available to date, I focus on examining contextual and temporal consistency of personality across multiple distinct daily contexts, using mainly naturalistic observations. By following the animals in their daily activities, I record several naturally occurring behaviours, previously shown to reflect four personality traits: Sociability, boldness, exploratiVeness, and anxiety. The chimpanzees' personality assessment was further explored by comparing different methods (i.e., ratings, experimental and naturalistic observations) to understand the contribution of each method for assessing the four personality traits. Firstly, I found evidence of both contextual and temporal consistency of personality-linked behaviours in chimpanzees and orangutans. The findings suggest that chimpanzee and orangutan personality can be reliably measured within ecologically relevant contexts and over time using naturalistic observations. Additionally, the results revealed different patterns of personality consistency across individuals, suggesting that some individuals may be more malleable than others in the expression of their personality-linked behaviours. Finally, I found that the three methods measured similar facets of targeted traits, highlighting the importance of developing methods that can capture the diversity of the expression of a personality trait. Overall, this thesis provides empirical evidence that a naturalistic method can measure personality consistency of both sanctuary chimpanzees and rehabilitant orangutans across distinct daily contexts and over time. By combining a naturalistic approach with different methods of assessment, we can better encompass the complexity of personality, in terms of how individuals express personality, how personality is measured or the factors that may influence its expression.

Content list

Declaration

Whilst registered as a candidate for the above degree, I have not been registered from any other research award. The results and conclusions embodied in this thesis are the work of the named candidate and have not been submitted for any other academic award.

Word count: 44, 258

Acknowledgements

First and foremost, I would like to thank Dr. Marina Davila-Ross who gave me the opportunity, 6 years ago, to work alongside her and observe primates in the field. Thank you for mentoring me all these years, for always encouraging me to think outside of the box (and think of the bigger picture) and trust my work, and finally for your tireless efforts to make me write more concisely (still not there fully but not far to go!). I would like to thank Prof. Kim Bard for her guidance, support, and (sometimes) challenging but always constructive discussion throughout my journey. I also want to thank Dr. Jérôme Micheletta for his support, advice (especially for not losing sight of the basics) and honest feedback regarding my work. Thank you all of you for being able to give me such valuable feedbacks on my dissertation in such a tight schedule in the last (mad) few weeks of my PhD, I will always be grateful for that!

I want to thank Dr. Leanne Proops and Dr. Marine Joly for their comments during my major and annual reviews, everyone from the admin and technician offices for all their help throughout my PhD, and I also want to thank the University of Portsmouth for the financial support.

I want to thank everyone in Chimfunshi Wildlife Orphanage, Zambia, who helped me during my research and made it all possible. More specifically, I would like to thank Sylvia Jones, Innocent Mulenga and Thalita Calvi for their hospitality and assistance while in Zambia. I would also like to thank Edwin van Leeuwen and Chimfunshi Research Advisory Board for their help and feedback. I am grateful to the founders of Chimfunshi, Sheila and David Siddle who built this wonderful place where these amazing chimpanzees can live freely for the rest of their lives. I would additionally like to thank all the staff from Chimfunshi, and particularity all the keepers (Tom, Mumba, Roger, Richard, Felix, Joseph and Robert) for their help during the identification of chimpanzees, their suggestions of improvement of my experimental stimuli (especially the snake!), their time for answering the personality questionnaires and sharing their culture (and teaching me how to cook nshima on lunch break!). Thanks to Jo-Anna Tilbury and Jane Copland for their advice and crucial help while dealing with visa and permit. Natotela sana!

The recordings from Sepilok Orangutan Rehabilitation Centre, Sabah, Malaysia, were obtained from Florence Rocque (Research permit numbers: JKM/MBS.1000-

2/2JLD.4 (7) / JKM/MBS.1000-2/2JLD.5 (48)). I would like to thank Augustine Tuuga, Sylvia Alsisto, Sailun Aris, and Sabah Wildlife Department and Sabah Biodiversity Centre for granting access to the orangutans, providing support, advice and hospitality to Florence throughout her PhD research in Malaysia. I am especially grateful to Florence Rocque for collecting all recordings of orangutans with the help of many people (Alice Judge, Macauly Gatenby, Fabio Crepaldi, Pravind, Yuen Zhao Yong, Jessie Hay, Aurelie Saulton, Trevor Dodds, Rezal) and staff members (Amoy, Victor, Bonny, Pamik, James, Jomius, Rufinah, Ibritinah, Dr Laura, Dr Pakee, Adrianus, Beddy, Beatrice). Thanks to all the rangers for their tireless efforts to take care of all lovely orangutans. Terima kasih banyak!

I want to thank all students and research assistants who helped me with data collection, video coding, and reliability coding during my PhD. I am grateful to Bruce Rawlings, Karri Neldner, Alex Burling, and Alex Lourie who I never met but collected the recordings from Chimfunshi in 2013, Berke Aydas for reliability coding, and Carla Osterburg as well as Koray Yilmaz for their amazing work with the video coding of the orangutans. While in Chimfunshi, I had the chance to work alongside researchers from other institutions who provided help with my (challenging) experiments, and who became my friends: thank you, Sarah De Troy, Manon Schweinfurth and Bianca Dietrich. More specifically, I would like to thank my amazing three research assistants – Sofia Vickers, Claire Barrault, Mackenzie Hewes – who helped me collect the recordings from Chimfunshi in 2017, create Monty, so I could test how the lovely chimps respond to a monitor lizard hidden in the bush (it is a shame it did not have many opportunities to shine!); thank you for making this fieldwork less challenging!

Of course, this research would not have been possible without the presence of the lovely chimpanzees and orangutans. I will always remember two particular individuals, the orangutan Matamis who I met before this project even started and the chimpanzee Chrissy who I met during this PhD. Two different individuals but one common way of living: Food is worth everything in this world!

Thank you to my colleagues and office-mates who survived being by my side during the past three years, and especially Derry and Zeynep for the support, puzzle sessions and laugh! Thank you, Flo (again!), for our brainstorming sessions while developing the coding scheme and always distracting me with your amazing stories!

I would like to thank all my friends who supported me throughout those years and were present when I needed them, thank you all! Thank you Stéphanie and Ophélie for our childhood friendship, for always believing in me from the beginning and encouraging me till the last minute. Eglantine, we met just before it all started but we went through this journey together; thank you for our friendship, for your valuable advice during the good and bad moments and for all your support (and especially the past few weeks!); we made it! Last but not least, I am deeply grateful to my family, especially my parents and my grandmother, who have always supported me and encouraged me to pursue my ambitions and dreams. Lorin, thank you for putting up with me, especially the last few months of this adventure (as you will remember, that was not easy!). Thank you for always being there when I needed it the most, and always believing in me, I am eternally grateful! Thank you, everyone, for riding this emotional roller-coaster with me!

Dissemination of research from this thesis

Publications, presentations, and posters

Chapter 3

- Chotard, H., Bard, K., Micheletta, J., and Davila-Ross, M. (2018). Semi-wild chimpanzees' personality measured using a naturalistic approach. Oral Presentation at the Primate Society of Great Britain (PSGB; April 11-12th)
- Chotard, H., Bard, K., Micheletta, J., and Davila-Ross, M. (2018). Measuring personality during naturally occurring situations in semi-wild chimpanzees (*Pan troglodytes*) living in a Zambian sanctuary. Oral Presentation as part of the Symposium *"*What contributions have sanctuaries made to our understanding of great ape cognition and behaviour? A compilation of latest findings", at the International Primatological Society (IPS; August 19th-25th)
- Chotard, H., Bard, K., Micheletta, J., and Davila-Ross, M. 3. Personality consistency measured using naturalistic observations across daily contexts in sanctuary chimpanzees (*Pan troglodytes*). *In preparation.*

Chapter 4

- Chotard, H., Bard, K., Micheletta, J., and Davila-Ross, M. (2018). Semi-wild chimpanzees' personality measured using a naturalistic approach. Oral Presentation at the Primate Society of Great Britain (PSGB; April 11-12th)
- Chotard, H., Bard, K., Micheletta, J., and Davila-Ross, M. Assessing personality in sanctuary chimpanzees (*Pan troglodytes*): Comparing naturalistic, rating and experimental approaches. *In preparation.*

Chapter 5

- Chotard, H., Bard, K., Micheletta, J., Rocque, F., Osterburg, C., Yilmaz, K., and Davila-Ross, M. Personality measured in daily contexts of rehabilitant Bornean orangutan (*Pongo pygmaeus*) using naturalistic observations. *In preparation.*

List of tables

List of figures

1. General Introduction

1.1. Overview

People's daily life is likely to be notably shaped by their personalities which can influence their decision making (Dahlbäck, 1990, 1991; Lauriola & Levin, 2001), and their relationship with others (Asendorpf & Wilpers, 1998; van Vianen & De Dreu, 2001). Within society, each person plays a specific role, a role that is defined by people's behaviours reflecting specific personality traits ("dimension of personality used to categorize people according to the degree to which they manifest a particular characteristic"; Burger, 1997), such as extraversion or anxiety. For instance, in a work environment, different leadership styles can emerge, where a different expression of extraversion, agreeableness, conscientiousness and openness to experience is expected (Judge & Bono, 2000; Walumbwa & Schaubroeck, 2009). Each individual's personality profile can be characterized, in general, by a certain combination of traits that are dynamically linked with one another, making each individual unique (Allport, 1961).

Because individuals can express their personality with multiple behaviours in multiple situations or contexts over time, personality remains a complex concept to fully examine and understand. For both human and nonhuman species, there have been multiple attempts to try to provide a simple and clear definition that would satisfy everyone (Gosling, 2001, 2008; John, Robins, & Pervin, 2008). Human personality has been defined by psychologists as "a dynamic organisation, inside the person, of psychophysical systems that create the person's characteristic patterns of behaviour, thoughts and feelings" (Allport, 1961) or as "those characteristics of individuals that describe and account for

consistent patterns of feeling, thinking and behaving" (Pervin & John, 1997). The current chapter is primarily about nonhuman primate personality. However, I will first review the literature about human personality research before focusing on the core topic of this thesis. By taking this step, I highlight what researchers can learn about human personality by studying nonhuman primates, and vice versa. Indeed, nonhuman primates are the closest relatives to humans (Perelman et al., 2011). By studying the expression of their personality-based behaviours, researchers can provide important insight into the understanding of how human personality emerged and possibly help identify the diverse selective pressures that might have acted across evolution (Buss, 1988; Nettle, 2006).

In humans, inter-individual behavioural differences were reported to be stable over time (Roberts & DelVecchio, 2000; Weisbuch, Slepian, Clarke, Ambady, & Veenstra-VanderWeele, 2010) as well as across different situations (Funder & Colvin, 1991; Furr & Funder, 2004; Moskowitz, 1982; Sherman, Nave, & Funder, 2010; Weisbuch et al., 2010). Yet, the latter was subjected to more debate views for over twenty years ("person-situation debate": Kenrick & Funder, 1988; see also Mischel, 1968; Mischel & Peake, 1982; Mischel, Shoda, & Mendoza-Denton, 2002). Based on Mischel's finding (1968), where a lack of stability across situations was reported, some psychologists claimed, for a long time, that behaviours were determined more by the situation than the person; i.e., the behaviour of an individual is fundamentally dependent on situational cues, resulting, thus, in a lack of consistency when measured across different situations (Mischel & Shoda, 1995; Mischel et al., 2002; Shoda, Mischel, & Wright, 1994). Since then, empirical research has provided much evidence to support the notion of a consistency across situations (Furr & Funder, 2004; Moskowitz, 1982; Weisbuch, Slepian, Clarke, Ambady, & Veenstra-VanderWeele, 2010), and it seems that higher consistency is found across

similar situations (Funder & Colvin, 1991). A situation is likely to influence people's response, but their individuality is most likely to be maintained across various situations.

People can be differentiated from one another based on visual cues such as gender, age and phenotypic characteristics (Bressan & Zucchi, 2009; Penn & Frommen, 2010) but also based on their personality profile. Personality defines, to some extent, who we are and is expressed by our behaviours. Whereas some individuals might, for instance, be more bad-tempered, others might be mellower in their everyday life. The way we behave in our daily life depends on what we go through on the moment as well as our past experience. Personality may be considered as a sort of social identity (Tajfel, 1974) contributing to giving an individual a place within society. An important aspect towards fully understanding individuals is to observe them interacting with both their physical and social environment in a range of situations; then, we are more likely to obtain a clearer overview of what their personality looks like as a whole.

Personality has been examined using different theoretical approaches. For instance, psychologists who used a biological approach argue that personality differences result from inherited characteristics, which are unarguably strongly linked to physiological processes (John, Robins, et al., 2008). Biological theorists believe that the nervous system that characterises each individual influences their behavioural responses to the environment (Eysenck, 1952, 2013). These theorists argue that this behavioural response is controlled by the interaction of two systems: behavioural activation and behavioural inhibition system (Gray & MacNaughton, 2003). Furthermore, trait theorists believe that individuals have a number of stable traits or specific characteristics that reflect their personality (Allport, 1961; Cattell, 1943). The trait theory suggests that the traits that

characterise one's personality lead the individuals to behave in a certain way and should remain stable over time and across situations (Winter & Barenbaum, 1999). These theorists are more interested in creating a comprehensive description of individuals' personality than understanding the causes of identified traits (Winter & Barenbaum, 1999).

A wide range of models were developed to describe human personality (John, Naumann, & Soto, 2008), but there is one that particularly prevailed in human personality research: the Five-Factor Model (Digman, 1990; Goldberg, 1993; McCrae & Costa, 1987, 1997). This model relies on a lexical approach which uses adjectives and descriptive phrases to describe differences between people's personalities (Digman, 1990). Human personality was described around five dimensions [i.e., broad domains encompassing a variety of related traits (Costa & McCrae, 1988)] – extraversion, agreeableness, conscientiousness, neuroticism and openness – and each of these dimensions was reported to show temporal stability (Costa & McCrae, 1988).

Compared to human personality research, it often seems notably more difficult to understand why animals behave, feel and think the way they do. Animal personality has already caught researchers' attention in the early 20th century. One of the first pioneers was Ivan Pavlov, who defined four types of personality in dogs: "excitable, lively, quiet and inhibited" (Pavlov, 1906). This work was followed by other comparative psychologists who started developing experimental and rating methods to measure individual differences in chimpanzee behaviours (Crawford, 1938; Hebb, 1949; Yerkes & Yerkes, 1936). In the 1980s, personality research experienced a significant increase of interest and focused on various species (for reviews, see Gosling, 2001; Gosling & John, 1999; Weinstein, Capitanio, & Gosling, 2008). As in humans, inter-individual differences in animals have

been reported to be stable over time and across situations/contexts, again with more attention put on the temporal consistency (Réale, Reader, Sol, McDougall, & Dingemanse, 2007; Sih, Bell, Johnson, & Ziemba, 2004). This lack of focus on contextual consistency in animal personality research will be addressed in the following section.

In this chapter, I review the current literature on personality in nonhuman species, with a special focus on nonhuman primates. I will provide an in-depth critical review of both temporal and contextual consistency in personality research with nonhuman primates. I discuss the background of the four personality traits of interest in my work (sociability, boldness, exploratiVeness, and anxiety) and present the reasons why they are relevant to study. I briefly discuss the potential effects of both intrinsic and extrinsic factors on personality expression, and I explain why I chose to study chimpanzees and orangutans as model species in the empirical chapters. Finally, this introductory chapter closes with an outline of my thesis.

1.2. Temporal and contextual consistency in nonhuman primates

Like humans, animals differ as individuals from one another in the way they consistently express their behaviours across multiple situations (or contexts) and different periods of time; this phenomenon refers to the definition of personality used in this thesis (Gosling, 2001, 2008). Time and context are, therefore, two key criteria to consider when examining both human and animal personality. Animal personality research has tested for such consistencies across a wide range of species, ranging from invertebrates (Chapman, Thain, Coughlin, & Hughes, 2011) to vertebrates (see Gosling, 2001 for a review; Vonk, Weiss, & Kuczaj, 2017), in order to be able to make claims that the studied animals have personality traits [i.e., enduring and quantifiable behavioural characteristics of an

individual that can be shared by other individuals of the same species (Carter, Feeney, Marshall, Cowlishaw & Heinsohn, 2013; Réale et al., 2007)]. Hence, the word personality is increasingly widely accepted in the animal literature. Other uses of the term include behavioural syndrome (Sih, Bell, Johnson, et al., 2004), temperament (Réale et al., 2007) or coping style (Koolhaas et al., 1999). In this section, I will present an overview of the literature that examined temporal and contextual consistency in nonhuman primates. By doing so, I provide the knowledge foundations that will help understand how this current doctoral project emerged. The papers used in this literature review are presented in Appendix 1, Table A.1).

Reviewing all empirical studies of the field of personality in nonhuman primates would have been a challenging task to undertake for the current project. Therefore, to establish this literature review, I tried to provide a broad overview of past research starting from the early 40s to the present day. I included empirical works that reported inter-individual differences in the behaviours of individuals in a wide range of nonhuman primate species. Studies that used ratings, experimental or naturalistic observations to measure personality were included in the overview and a particular focus was given to studies that examined systematically temporal and contextual consistency of personality. A total of 186 studies was included in this literature review. The overview of these studies presented in Table A.1 provides extensive information about species studied, demographics (age, sex), setting (captivity, semi-wild, wild), method of assessment, context examined when behavioural observations were carried out, personality traits measured, results about consistency (temporal, contextual), and additional results (sex and age differences, method comparisons).

Among these 186 nonhuman primate studies, 59 studies examined consistency in personality. The majority of these 59 works (70%) found temporal consistency, which was reported in 21 different primate species (Dammhahn & Almeling, 2012; Ebenau, von Borell, Penke, Ostner, & Schülke, 2019; Tomassetti et al., 2019; Uher, Asendorpf, & Call, 2008). Both short (e.g., a few weeks apart: Tomassetti et al., 2019; Uher et al., 2008) and long-term (e.g., years apart: Koski, 2011b; Seyfarth, Silk, & Cheney, 2012) temporal consistency were found in primates. Temporal consistency has been demonstrated with the rating method (Stevenson-Hinde, Stillwell-Barnes, & Zunz, 1980b; Weiss et al., 2017), experimental observations (Massen, Antonides, Arnold, Bionda, & Koski, 2013; Šlipogor, Gunhold-de Oliveira, Tadić, Massen, & Bugnyar, 2016) and naturalistic observations (Koski, 2011b; von Borell, Kulik, & Widdig, 2016)[1]. Temporal consistency of behaviours or personality traits certainly suggests that they are an enduring characteristic of individuals' personality structure. Yet, if one wants to examine and fully understand the personality structure of a species, it is important to examine both temporal and contextual consistency.

Among the 59 studies that reported consistency, only 24% systematically examined the contextual consistency of personality (see Table A.1), which was reported for 10 different primate species. To date, consistency across contexts was only tested using experimental observations (Carter, Marshall, Heinsohn, & Cowlishaw, 2012b; Dammhahn & Almeling, 2012; Massen et al., 2013; Šlipogor et al., 2016; Uher et al., 2008; Uher,

[1] A review of the different methods used to assess personality in nonhuman primates is provided in Chapter 2.

Werner, & Gosselt, 2013). Although this past research contributed greatly to the field by demonstrating that individual nonhuman primates can be consistent in their personality-related behaviours across situations, these studies examined consistency of personality traits across very specific situations (e.g., small vs big novel objects to measure exploratiVeness). Therefore, what the field of personality requires now is to assess systematically personality consistency across broader contexts.

From a general point of view, personality is a complex concept to study and its complexity can be conceptualised on three levels: behaviour, context and time. Individuals can express their personality traits through diverse behaviours – e.g., sociability may be measured when an individual shares food (Silk, Brosnan, Henrich, Lambeth, & Shapiro, 2013) or sits in proximity of others (Massen & Koski, 2014). In addition, individuals may express their personality in various contexts and throughout their lives. A context can be defined as a broad category that regroups multiple specific situations (Sih, Bell, & Johnson, 2004). For instance, a feeding context can be both with and without predators (Sih, Bell, & Johnson, 2004) and vigilance context can be related to both food anticipation (Waitt & Buchanan-Smith, 2001) or background disturbance (Kutsukake, 2006; Shultz, Faurie, & Noë, 2003). These contexts can vary on different levels such as functional (e.g., feeding vs resting), affective (e.g., play vs aggression), arousal (e.g., vigilance vs grooming) as well as social (vs solitary). Consequently, the context seems to play an important role in the understanding of a personality trait where multiple facets of expression – i.e., large set of specific and related aspects of a trait; (Costa & McCrae, 1988) – of the same trait can be expected – e.g., curiosity towards an object (Uher et al., 2008) or curiosity in an environment (Watson & Ward, 1996). Thus, if we want to fully

grasp the diversity of expression of different personality traits, we need to measure the behaviours of animals on multiple occasions and in multiple contexts.

Capturing such a diversity of expression of personality traits may be achieved by using naturalistic observations (Gosling, 2008; Mehta & Gosling, 2008; Uher, 2008a). However, using this method to measure systematically temporal and contextual consistency is much ignored in nonhuman primate personality research in comparison with experimental observations (see Table A.1). Naturalistic observations may provide a comprehensive representation of the personality traits targeted as multiple behaviours displayed in multiple contexts and occasions are likely to be covered. Combining such an approach with ratings and/or experimental observations could contribute to an improved understanding of the personality of nonhuman primate species (see Chapter 2 for a review).

Following the definition of personality, individuals are expected to differ from one another in the way they express their behaviours but should show some consistency when observed on different time periods and across different contexts/situations (Gosling, 2001; Vaugh et al., 2017). Researchers consider personality consistency when individuals show stability in their personality-based behaviours while maintaining their differences with other members of the group, i.e., the rank-order consistency [stability of the relative position of individuals on a personality trait over time and context; (Roberts & DelVecchio, 2000)]. Although consistency is a key aspect of personality, ecologists investigating animal personality also paid attention to the plasticity aspect of personality (Dingemanse, Kazem, Réale, & Wright, 2010; Dingemanse & Wolf, 2013; Wolf & Weissing, 2012). Individuals may differ in the way they respond to the environment where different patterns of temporal and contextual consistency in behaviours may be observed

across individuals. Dingemanse and colleagues (2010) developed a framework to study behavioural consistency and plasticity in animals, i.e., the behavioural reaction norm. This approach incorporates the range of behavioural phenotypes that an individual produces in a set of environmental conditions, allowing researchers to jointly examine animal personality and individual plasticity within the same framework. While such a concept is applied in the field of ecology (e.g., Araya-Ajoy & Dingemanse, 2017; Carter, Goldizen & Heinsohn, 2012; Carter et al., 2013; Dingemanse et al., 2012), it has not received much attention in psychological research that examined nonhuman primate personality (Suomi, Novak, & Well, 1996; Uher, Addessi, & Visalberghi, 2013; Uher et al., 2008).

To my knowledge, only Uher and colleagues (2008, 2013) and Suomi and colleagues (1996) conducted individual-oriented analyses to investigate the stability of the individuals' characteristics. Some individuals may show high internal consistency in their behaviour throughout their lives in order to increase their individual fitness (reproductive success: Réale, Gallant, Leblanc, & Festa-Bianchet, 2000; life span: Altschul et al., 2018). Such a pattern of consistency might be particularly important for social animals as predictability is considered a key element to ensure group cohesion (Webster & Ward, 2011). However, individuals may also show behavioural variations to increase their benefits in terms of fitness (Wolf & Weissing, 2012). For some animals, personality can be malleable and sensitive to changes, allowing individuals to adapt and cope with new situations (Dingemanse & Wolf, 2013; Koolhaas et al., 1999). Nonhuman primates are known to be capable of social learning (Whiten, 2000), where they acquire new behaviours and skills throughout their lives (Jaeggi et al., 2010; Matsuzawa et al., 2001). For instance, Sapolsky and Share (2004) reported in a troop of olive baboons (*Papio anubis*) that individuals immigrating into a new group adapted their personality type to the members of

the group. Such changes in behaviours are also observable in human society where people changing countries may need to adjust some of their behaviours to the new culture in order to increase their chance of successful integration (Berry, 1997, 2005). Thus, both consistency and malleability in the behavioural expression of a personality trait depending on the context and time period might be observable and potentially valuable in terms of fitness (Jarrett, Bonnell, Young, Barrett, & Henzi, 2018).

To summarize, this critical review of research on nonhuman primate personality consistency emphasised the main gaps in our current knowledge regarding both temporal and contextual consistency in nonhuman primate personality research. Based on this information, the three empirical chapters of this thesis will address the gap in the literature and focus particularly on the contexts.

1.3.Personality traits of interest

Personality traits may be classified as either universal trait dimensions or species-specific trait dimensions (Uher, 2008a, 2008b). The universal traits are expressed in different species allowing cross-species comparisons. If the mean and variance of the trait distribution are comparable across the species, then we call them weak universals and direct comparisons of these traits between individuals of different species are possible. If the mean and variance of the trait distribution are significantly different across species, they are considered as strong universals; direct comparisons are also possible, provided that the trait scores are standardized. In contrast, the species-specific traits are only expressed by the individuals of a given species (Uher, 2008a, 2008b).

All behaviours that may reflect the same personality trait have to be considered when assessing individuals' personality, which would help cover multiple facets of expression of a personality trait. To achieve this step, it is important to take into consideration the ecology and the behavioural repertoire of a given species (Uher, 2008a), allowing the identification of relevant behaviours (i.e., universal and species-specific behaviours) that may reflect a personality trait.

It is essential to try to capture all behaviours that index personality traits and to examine them with the different methods applied in personality research (ratings, experimental and naturalistic observations). Each method of assessment has its own limitations (see Chapter 2 for a review), so using one method in combination with another is likely to complement each other assessment. Such a detailed evaluation is likely to help provide a comprehensive description of the personality profile of individuals (Carter, Marshall, Heinsohn, & Cowlishaw, 2012a).

Furthermore, most of the previous research on nonhuman primate personality (about 59% of the reviewed empirical studies; see Table A.1) focused on boldness and exploratiVeness (or at least some aspects of these personality traits). Subsequently, it is important to include more of the other personality traits in nonhuman primate research, which would help to provide a broader picture of how the personality profile of a species is structured and how these different traits are entwined with one another (Koski, 2014).

In the current thesis, four personality traits were investigated: sociability, boldness, exploratiVeness and anxiety. These four traits have been both reliably rated (Clay, Bloomsmith, Bard, Maple, & Marr, 2015; Freeman et al., 2013; Weiss & King, 2015) and measured behaviourally (Anestis, 2005; Baker & Aureli, 1997; Koski, 2011b; Massen et

al., 2013; Uher, Asendorpf, & Call, 2008) by human observers. In addition, these traits are ecologically and evolutionary relevant (Réale et al., 2007; B. R. Smith & Blumstein, 2008) as they are expressed in different primate species suggesting their universality (Uher, 2008a) and are argued to affect individual fitness (e.g., life span: Altschul et al., 2018; survival: Silk, Alberts, & Altmann, 2003).

Regarding the selection of the behaviours to reflect the four personality traits of interest, an explorative approach was used as I wanted to capture a relatively broad picture of the different personality traits. I tried to target any behaviours that had, to some extent, some connotations with the targeted trait. I considered a wide range of studies in nonhuman primates (see Table A.1), including personality research but also behavioural research in which personality assessment was not necessarily the main aim of the study. Therefore, some of the chosen behaviours might not be necessarily strong and reliable indicators of the personality trait of interest (e.g., food sharing representing sociability). The choice of the behaviours thought to reflect each of these four personality traits is discussed below in relation to the literature.

1.3.1. Sociability

Sociability is defined as a behavioural reaction of an individual towards the presence or absence of conspecifics and the closeness to others (Réale et al., 2007). Previous research reported that some personality traits may be divided into sub-traits (Carter et al., 2012a; Koski, 2011b; Tkaczynski et al., 2018). Indeed, it has been suggested that sociability could be characterised by multiple sub-traits such as tactility, equitability, and positive affect (Koski, 2011b; Tkaczynski et al., 2018). Sociability can be measured in the daily life of the individuals through their behaviours as their proximity with others (Koski, 2011b), food

sharing (Silk et al., 2013), food begging (Freeman et al., 2013), body contact (Ebenau et al., 2019), grooming (Eckardt et al., 2015; Neumann, Agil, Widdig, & Engelhardt, 2013) or play (Koski, 2011b; Suomi et al., 1996).

Sociability seems to have important ecological and evolutionary consequences in individuals‘ fitness (Wolf & Weissing, 2012). For animals living in social groups, their level of sociability is likely to determine their position within their group and, therefore, affect their future (Perry, 1998). Past research showed that social individuals tend to have a higher reproductive success (Parish, 1996) and a higher survival rate regarding themselves (Archie, Tung, Clark, Altmann, & Alberts, 2014), but also their offspring (Silk et al., 2003, 2009) in comparison with less social individuals.

1.3.2. Boldness

Boldness represents the tendency to engage in risky situations (Réale et al., 2007). Boldness has been documented in various animal species suggesting its universality across taxa (Koski, 2014; Réale et al., 2007; Sih & Bell, 2008). This personality trait can be to some extent reflected by various behaviours such as hitting, chasing or biting conspecifics (Clay et al., 2015; Freeman et al., 2013), as well as approaching threatening objects (Dammhahn & Almeling, 2012; Koski & Burkart, 2015; Santillán-Doherty et al., 2010).

Boldness was reported to have fitness consequences in humans and animals (B. R. Smith & Blumstein, 2008; Wilson, Clark, Coleman, & Dearstyne, 1994). The different levels of risk-taking of individuals may lead to different outcomes (Dingemanse & Réale, 2005; B. R. Smith & Blumstein, 2008). By taking more risks, some individuals may, for instance, increase their chances to mate with multiple partners, and, consequently, increase

their reproductive success (Réale et al., 2000a; Reaney & Backwell, 2007). However, by taking more risks, individuals may also increase their chances to encounter predators, thus leading to a shorter life span (Bremner-Harrison, Prodohl, & Elwood, 2004; Dugatkin, 1992, 2013).

1.3.3. Explorativeness

The explorativeness of individuals is characterized by their interest directed towards their social and physical environments or their behavioural reactions towards a new situation (Réale et al., 2007). Explorativeness has been documented in various species suggesting its universality (Gosling, 2001; Réale et al., 2007). This personality trait is arguably readily observable through various behaviours such as exploring a new environment (Dammhahn, 2012; Watson & Ward, 1996), gazing (Forss, Schuppli, Haiden, Zweifel, & van Schaik, 2015; Schuppli et al., 2017), approaching (Massen et al., 2013; Santillán-Doherty et al., 2010), or manipulativeness (touching or handling an object) (Damerius, Graber, Willems, & van Schaik, 2017; Uher et al., 2008).

Explorative individuals seem to have a higher survival chance, but to a small degree (Dingemanse, Both, Drent, & Tinbergen, 2004; Réale et al., 2007; B. R. Smith & Blumstein, 2008). It could be inferred that showing high rate of explorative behaviours may increase the chance of individuals (especially young individuals) to learn new skills (Schuppli, Forss, et al., 2016; Schuppli et al., 2017; Schuppli, Meulman, et al., 2016; van Noordwijk & van Schaik, 2005), which in return can possibly help them survive in adulthood (Russon, 2006; van Adrichem, Utami, Wich, van Hooff, & Sterck, 2006).

It is important to note that exploratiVeness and boldness are sometimes difficult to separate as a novelty can also be threatening (Carter, Feeney, Marshall, Cowlishaw, & Heinsohn, 2013; Coleman & Wilson, 1998). For instance, previous research used similar experimental assays to measure these two personality traits such as novel object tests [boldness: (Dammhahn, 2012); exploratiVeness: (Uher et al., 2008)]. Using multiple experimental assays and measuring multiple behaviours to assess the same personality trait are likely to avoid confusion (Carter et al., 2013).

1.3.4. Anxiety

This personality trait represents fearful responses to potential dangers (Boissy, 1995). Anxiety can be measured using various behavioural indicators such as yawning (Baker & Aureli, 1997), self-directed behaviours (self-touch, self-scratch) (Aureli & Waal, 1997; Baker & Aureli, 1997), vigilant (Kalin & Shelton, 2003; Kutsukake, 2003), rocking or fleeing behaviours (Clay et al., 2015; Uher et al., 2008).

Anxiety is likely to have important consequences on an individual's fitness. For instance, being vigilant to the surrounding by scanning for predators (Elgar, 1989; Quenette, 1990), may increase an individual's survival chances as the individual may react more quickly and according to the situation. However, showing a high level of anxiety can also have long-term consequences on health (Maestripieri & Hoffman, 2011; Mendoza, Capitanio, & Mason, 2000).

1.4.Effect of intrinsic and extrinsic factors on personality

There are many factors that can influence the likelihood of expressing a personality trait, such as intrinsic (age, sex, genetics) or extrinsic (environment) factors. To better

understand why an individual behaves in a certain way that differs from the others, it is important to consider these factors and try to determine to what extent they can influence the expression of a personality trait.

1.4.1. Intrinsic factors: Demographic variables

Age

Past research in nonhuman primates showed that the expression of some personality traits can differ between age groups where, for instance, younger individuals tend to be bolder, more active and more social than adults (Baker, Lea, & Melfi, 2015; Carter, Marshall, Heinsohn, & Cowlishaw, 2014; King, Weiss, & Sisco, 2008; Massen et al., 2013). Age differences are often explained in relation with learning abilities. For instance, being bolder, more explorative and playful allows younger individuals to develop motor skills (Fagen, 1981; Spinka, Newberry, & Bekoff, 2001), giving them the opportunity to learn new survival skills (Russon, 2006; van Adrichem et al., 2006) or to develop social abilities (Palagi & Paoli, 2007; Pellis & Iwaniuk, 2000; Poirier, Bellisari, & Haines, 1978) which can have an impact on individual fitness (Silk et al., 2003).

However, no studies tested whether the age of individuals could affect the consistency of personality, or more specifically, contextual consistency. Yet, it is possible that different patterns of consistency in personality emerge as individuals face different challenges, experience, and learn new skills throughout their life (Russon, 2006; van Adrichem et al., 2006; van Noordwijk & van Schaik, 2005). Their personality traits are likely to be shaped by various experiences and, therefore, different patterns of consistency may be expected across the development. Early in their lives, younger individuals learn

new behaviours or skills that will ultimately help them survive into adulthood (Russon, 2006), but because they are still learning these skills, they may show more malleability in their personality-linked behaviours when expressed across contexts. In contrast, older individuals may benefit from more consistency in their behaviours and decisions to facilitate their interactions with both their physical and social environments. Yet, both young and old individuals may need to fine-tune their behaviours to the situation they are in (Allport, 1961; Dingemanse & Wolf, 2013), which is likely to result in different patterns of consistency for different personality traits, allowing the individuals to increase their fitness (Wolf & Weissing, 2012).

Sex

Previous research in primates revealed conflicting findings regarding the effect of sex on the personality score. Whereas some studies showed that males (relative to females) were more anxious, active, dominant and less social (Koski, 2011b; Manson & Perry, 2013; J. S. Martin & Suarez, 2017; McGuire, Raleigh, & Pollack, 1994; Pederson, King, & Landau, 2005), other studies did not find any sex differences (Carter et al., 2014; Damerius, Forss, et al., 2017; Dammhahn & Almeling, 2012; Šlipogor et al., 2016). These findings could possibly result from differences in the method used to assess personality, the group composition or the ecology of the species; all these factors are likely to differ across studies.

There is a wide range of social organizations across nonhuman primate species (Smuts, Cheney, Seyfarth, & Wrangham, 2008). This diversity of social structure is likely to have an impact on the expression of personality-based behaviours between males and females. For instance, male chimpanzees are known to form enduring social bonds with other males

to acquire and maintain their dominance rank (Mitani, 2009; Mitani & Watts, 2001); thus, males are likely to show higher level of social personality traits (e.g., equitability) than females (Koski, 2011b).

1.4.2. Intrinsic factor: Genetic variables

In order to better understand the behavioural basis of personality, researchers focused on the potential genetic influences on personality traits in various animal species (van Oers, de Jong, van Noordwijk, Kempenaers, & Drent, 2005; van Oers & Sinn, 2013). Genes are biological units that transmit characteristics from one generation to the other (Cooper, 2000), hence we should assume that behavioural characteristics are partly heritable (van Oers et al., 2005). Rather than being determined by one single gene, personality is likely to be determined by the actions of multiple genes influencing the expression of different traits (Barr et al., 2004; Hopkins, Donaldson, & Young, 2012; Latzman, Freeman, Schapiro, & Hopkins, 2015).

Higher heritability (i.e., relative amount of genetic variation in relation to the phenotypic variation) suggests a greater contribution of genetics in the variation of a personality trait. Such observations have been documented in various primates species, where various personality traits have been reported to be highly heritable, such as social impulsiveness in vervet monkeys (Fairbanks, Newman, et al., 2004) or dominance (Latzman, Freeman, et al., 2015; Weiss, King, & Enns, 2002; Weiss, King, & Figueredo, 2000) and extraversion (Latzman, Freeman, et al., 2015) in chimpanzees. However, it is important to remember that personality is not solely determined by genes, but also by environmental influences (Clarke & Boinski, 1995; Uher, 2011b).

1.4.3. Extrinsic factor: Environment

The idea that personality of individuals is consistent over time and across contexts is widely accepted in the field (Réale et al., 2007); yet, it is important to point out that the expression of personality can be also influenced by environmental factors (McDougall, Réale, Sol, & Reader, 2006). Individuals are likely to show some plasticity in their behaviour-based traits throughout their lives allowing them to adjust to the situation experienced and increase their fitness (Dingemanse & Wolf, 2013; Wolf & Weissing, 2012). Nonhuman primates, especially great apes, are known to show a certain flexibility in their behaviours (Davila-Ross, Allcock, Thomas, & Bard, 2011; Masi et al., 2012; Robbins et al., 2016; van Schaik, 2013; van Schaik et al., 2016; S. K. Watson et al., 2015). This plasticity in behaviours can result from ecological (Gruber et al., 2019; McLennan, Spagnoletti, & Hockings, 2017) as well as social (Amici, Call, Watzek, Brosnan, & Aureli, 2018; Flack, Jeannotte, & de Waal, 2004) influences.

Nonhuman primates live in diverse social organisations (Smuts et al., 2008), where different types of social interactions between individuals are likely to occur, such as affiliative (e.g., Mitani, 2009) or agonistic (e.g., Marzec et al., 2016), and influence the behavioural expression of individuals (e.g., Flack et al., 2004). To date, only a few empirical findings in the field of personality showed that the expression of personality traits differ across groups (Bard & Gardner, 1996; Koski, 2011b; Koski & Burkart, 2015; Šlipogor et al., 2016) and that closely bonded individuals tend to share similarities in some of their personality traits (Massen & Koski, 2014). This series of studies in nonhuman primate species suggest that personality can be shaped by the social environment leading, therefore, to different expression of personality traits, such as sociability-related traits

(Bard & Gardner, 1996; Koski, 2011b), exploration and activity (Bard & Gardner, 1996) in chimpanzees, as well as boldness (Koski & Burkart, 2015) and exploration (Koski & Burkart, 2015; Šlipogor et al., 2016) in common marmosets (*Callithrix jacchus*).

Nonetheless, disentangling ecological influences from social influences may be sometimes difficult. For instance, Koski (2011b) found cross-colony differences for different personality traits in zoo chimpanzees. Although important insight about the plasticity of personality in great apes was pointed out, the ecological influence resulting from different zoo environments could not be ruled out. It seems clear that individuals' behaviours can be influenced by the environment, but, how can one be sure which factors have more impact on the expression of personality? If one wants to determine how social influences could shape personality expression in nonhuman primates, it is important to study animals who have similar genetic background and live in similar ecological environment.

1.5.Examining personality in great ape species

Both humans (Kaplan, Hooper, & Gurven, 2009) and nonhuman primates (Smuts, Cheney, Seyfarth, & Wrangham, 2008) live in diverse social structures; this complex social organisation in humans might have led to the development of different personality types (Triandis, 2001). To understand how this diversity of personality profiles emerged in human society, nonhuman primates seem to be the ideal model. Research on personality traits, which are measurable through behaviours occurring in day-to-day contexts, in nonhuman primates could provide an important contribution to better understand the evolutionary history of human personality (Buss, 1988). More particularly, studying great apes, the closest relatives to humans (Goodman et al., 1998; Stauffer, Walker, Ryder,

Lyons-Weiler, & Hedges, 2001), in their natural environment would allow us to access a wide range of behaviours, some of which are commonly expressed across different species. By doing so, we can possibly highlight similarities or variations, which in return can help identify what selective pressures may have played a role in the emergence of some personality traits in humans (Buss, 1988; Nettle, 2006).

1.5.1. Chimpanzees (Pan troglodytes)

Chimpanzees are our phylogenetically closest relatives (together with bonobos), sharing a common ancestor approximately 4-6 million years ago and sharing 98% of our genome (Stauffer et al., 2001; Waterson, Lander, Wilson, & The Chimpanzee Sequencing and Analysis Consortium, 2005). Wild chimpanzees inhabit tropical forests and savannas of Equatorial Africa and are split into four sub-species: Western chimpanzee (*Pan troglodytes verus*), Nigeria-Cameroon chimpanzee (*P. t.ellioti*), Central chimpanzee (*P. t.troglodytes*), and Eastern chimpanzee (*P. t.schweinfurthii*). The species lives in multi-male, multi-female groups with typical fusion-fission structure, including between 20 to 150 individuals within the group (Boesch, Boesch, & Boesch-Achermann, 2000; Goodall, 1986; Nishida, 1990; Watts, 1998). Chimpanzees live approximatively 40-50 years in the wild (Boesch et al., 2000; Nishida, 1990) but can live longer in captivity. Sexual maturity is reached around 11 years of age and first reproduction occurs around 13 years of age (Boesch et al., 2000); the interbirth interval is approximately 5-6 years (Boesch et al., 2000). The species reproduce throughout the year and the gestation duration is about 8 months (Boesch et al., 2000; Humle, Maisels, Oates, Plumptre, & Williamson, 2016; Nishida, 1990). While males stay in their natal groups, the females emigrate to other groups during adolescence (Boesch et al., 2000).

Chimpanzees live in a hierarchical social group, where competition and cooperation between individuals set the group dynamic (Mitani, Watts, & Muller, 2002; Muller & Mitani, 2005). The males have a linear dominance structure whereas mixed results have been reported regarding the dominance structure for the females (Goodall, 1986; Lehmann & Boesch, 2009; Muller & Mitani, 2005; Nishida, 1990; Watts, 1998). Living in a competitive highly social environment where individuals compete to access food or females is likely to affect the individual fitness (Muller and Mitani 2005) and may eventually influence the expression of some behaviours, hence, personality traits.

Past research has focused on captive chimpanzees, and established both temporal and contextual consistency of personality traits (King & Figueredo, 1997; Koski, 2011b; Kutsukake et al., 2012; Massen et al., 2013). The rating approach described chimpanzee's personality by six dimensions - dominance, extraversion, conscientiousness, agreeableness, neuroticism, and openness (King & Figueredo, 1997; Weiss et al., 2017) – which, except for dominance, are referred to as the Big-Five in human personality research (Costa & McCrae, 1992). Both naturalistic and experimental approaches found similar personality structure but used a different classification (Réale et al., 2007). For instance, exploration tendency or curiosity can be perceived as an equivalent for openness (Massen et al., 2013; Uher & Asendorpf, 2008), and sociability as an analogue label for extraversion and agreeableness (Koski, 2011b).

1.5.2. Orangutans (Pongo pygmaeus)

Unlike chimpanzees, orangutans are our most phylogenetically distanced great ape relatives, sharing a common ancestor approximately 9-13 million years ago and sharing 97% of our genomes (Hobolth, Dutheil, Hawks, Schierup, & Mailund, 2011; Locke et al.,

2011). Wild orangutans inhabit tropical rainforest in Indonesia and Malaysia and are split into three species: Bornean orangutan (*P. pygmaeus*, with three subspecies), Sumatran orangutan (*P. abelii*), and Tapanuli orangutan (*P. tapanuliensis*). The species have a rather flexible social organisation; they are mostly semi-solitary but they are also known to form occasional social parties (Delgado & van Schaik, 2000; van Schaik, 1999). Throughout their lives, orangutans experience various social interactions, such as sexual encounters (Galdikas, 1985, 1995), mother-infant interactions (van Noordwijk & van Schaik, 2005) as well as occasional interactions between females (van Noordwijk et al., 2012). Orangutans live approximatively 50 years in the wild (Wich et al., 2004); and captive individuals can live longer. They reach sexual maturity between 10 and 15 years of age and the females tend to have their first reproduction at the age of 15 years (Wich, Atmoko, Setia, & Schaik, 2010); the interbirth interval ranges from 6 to 9 years (Wich et al., 2010, 2004). The species reproduce throughout the year and the gestation duration is about 8-9 months (Shumaker, Wich, & Perkins, 2008). During adolescence, females tend to form small home ranges which overlap to some extent with those of their mothers (Singleton et al., 2009). In contrast, males disperse further away from their mothers and enter in a transient phase where they do not have a fixed home range before settling down; males' home range often overlaps with that of several other males and multiple females (Utami Atmoko et al., 2009).

The personality structure has been primarily described using rating and experimental methods for captive orangutans (Forss et al., 2015; Uher et al., 2008; Weiss, King, & Perkins, 2006). Using the rating approach, five personality dimensions have been determined: extraversion, dominance, neuroticism, agreeableness and intellect (Weiss, Inoue-Murayama, King, Adams, & Matsuzawa, 2012; Weiss et al., 2006; Weiss & King,

2015). Additionally, Uher and colleagues (Uher & Asendorpf, 2008; Uher et al., 2008) who combined both experiments and ratings described the personality of the species in terms of 17 personality traits (e.g., aggressiveness, anxiousness, curiosity, friendliness to humans). Using experimental or naturalistic observations, other studies focused particularly on exploratiVeness in relation with cognitive abilities or social learning (Damerius, Forss, et al., 2017; Damerius, Graber, et al., 2017; Forss, Willems, Call, & van Schaik, 2016; Schuppli et al., 2017) allowing them to determine whether some rehabilitant orangutans were more successful after releases in the wild. However, none of these studies systematically assessed the contextual and temporal consistency of exploratiVeness.

1.6. Outline of the chapters in this thesis

In this PhD, I aim to measure the personality of chimpanzees and orangutans, the great ape species that are closest and most distant related to humans, respectively, across multiple distinct daily contexts using naturalistic observations. There is substantial evidence that both species' personality can be reliably measured using different methods of assessment (i.e., ratings, experimental and naturalistic observations). However, past research mostly focused on assessing personality consistency over time. Therefore, I specifically focus on examining contextual consistency of personality. Furthermore, there is limited research in regard to comparing systematically different methods to assess the same personality trait in nonhuman primates. In particular, I am interested in determining what different methods of assessment can tell us about the differential expression of a specific personality trait.

In Chapter 2, I provide some background information regarding the methodology used in nonhuman primate personality research. More specifically, I critically evaluate the three different methods (ratings, experimental and naturalistic observations) that have been

used to measure nonhuman primate personality, discussing the strengths and weaknesses of each. I discuss the different statistical methods used in the literature to estimate the reliability of the personality measurements. I will particularly focus on the two most commonly used statistical methods (Intraclass Correlation Coefficient and Cronbach's alpha) and a less common statistical method (Generalizability Theory) which will be used throughout this research. Finally, I conclude this chapter by presenting the two different approaches (i.e., traditional and conceptual approaches) to determine the personality structure of individuals.

In Chapter 3, I look at the behavioural consistency of 22 sanctuary chimpanzees across multiple distinct contexts and over two time periods (4 years apart) using a very detailed coding scheme that includes eight ecologically relevant contexts and 22 behaviours reflecting sociability, boldness, explorativeness and anxiety. The aim of this study is to test whether the personality-based behaviours of sanctuary chimpanzees are consistent across multiple distinct daily contexts in natural settings. Based on previous experimental studies in captive chimpanzees (Kutsukake et al., 2012; Massen et al., 2013; Uher et al., 2008), I hypothesise that the behaviours are consistent across naturally occurring contexts. Additionally, like previous research in captive chimpanzees (Koski, 2011b; Uher et al., 2008), I tested if there was a temporal consistency across two time periods. Finding contextual and temporal consistency in personality-linked behaviours would confirm the effectiveness of the naturalistic approach to highlight inter-individual differences in naturally occurring contexts.

In Chapter 4, I compare the three methods of personality assessment (i.e., ratings, experimental and naturalistic observations) in sanctuary chimpanzees measuring the same

personality traits (i.e., sociability, boldness, explorativeness, anxiety). A questionnaire that combined ratings of adjective-behaviour items to measure the four traits was developed, and two experiments to measure two of these traits, boldness and explorativeness, were conducted. These two sets of data were compared with the data obtained using the naturalistic approach in Chapter 3. In this study, I aim to determine whether each method measures the same trait, more specifically the same facets of expression of the personality trait. Based on past research (Carter et al., 2012a; Tkaczynski et al., 2018; Uher et al., 2008), I hypothesise that the three approaches show some similarities in their assessment of the personality traits. Finding similarities across the methods would provide a finer description and a better understanding of the complexity of the personality structure of a nonhuman primate species.

Previous studies in orangutans defined the personality of the species using primarily ratings and experimental observations. So far, there is no evidence that personality consistency can be measured using naturalistic observations. In Chapter 5, applying the same naturalistic approach developed in Chapter 3, I measure the behavioural consistency of 20 juvenile, adolescent and adult rehabilitant orangutans. Seventeen behaviours – reflecting sociability, boldness, explorativeness and anxiety – displayed in seven distinct daily contexts over two time periods (5 months apart) were measured. The aim of this study is to test whether orangutan personality-linked behaviours may be consistent across daily contexts and time periods. Based on the sanctuary chimpanzee study (Chapter 3) and the previous personality studies (Massen et al., 2013; Uher et al., 2008), I hypothesise that behaviours linked to personality traits in orangutans may show consistency across multiple distinct daily contexts and over time. In addition, I compare three groups of rehabilitant orangutans that differ in age and socioecological environment.

On a preliminary basis, I wanted to examine to what extent these three groups might show different patterns of contextual consistency in their personality-linked behaviours.

Finally, in Chapter 6, I provide an overview of the key findings of my three empirical studies and the implications in relation to the current animal personality literature. I present the major strengths (naturalistic observations, multiple contexts, method comparisons, and study of two great ape species) as well as limitations of my current research, and suggestions for future research. Overall, this thesis provides empirical evidence that a naturalistic method can measure personality consistency of both sanctuary chimpanzees and rehabilitant orangutans across distinct daily contexts and over time. Combining such a comprehensive approach with different methods of assessment may contribute to a better understanding and a detailed description of personality traits. We can then better encompass the complexity of personality, in terms of how individuals express personality, how personality is measured or the factors that may influence its expression.

2. Review of the methodology used in nonhuman primate personality research

Animals differ from one another in the way they express their personality through their behaviours. These inter-individual differences have been shown to be consistent across different situations or contexts throughout life (see section 1.2; Gosling, 2008; Réale et al., 2007). Different methods have been used in the field to assess this diversity of expression, to measure the contextual and temporal consistency aspect of personality, and to determine the personality profile of the animals.

Some of the reviews of the field of animal personality focused on all animal species (e.g., Carere & Maestripieri, 2013; Gosling, 2001, 2008; Koski, 2011a; Vonk, Weiss, & Kuczaj, 2017; Vazire, Gosling, Dickey, & Schapiro, 2007; Weinstein, Capitanio, & Gosling, 2008) and others focused particularly on nonhuman primates (Freeman & Gosling, 2010; Freeman, Gosling, & Schapiro, 2011; Itoh, 2002; Uher, 2011b; Weiss, King, & Murray, 2011). The review of the methodologies presented in this chapter relies on these earlier reviews and provides an up-to-date overview of the methodologies used in research on nonhuman primate personality (see Table A.1).

One hundred and eighty-six empirical studies on nonhuman primates were identified and reviewed here. Among those, 46% used the rating method to measure personality in 31 different primate species and only 24% of them tested and found temporal consistency of the measurement (King & Figueredo, 1997; Konečná, Weiss, Lhota, & Wallner, 2012; Manson & Perry, 2013; McGuire et al., 1994); to date, contextual consistency has never been assessed using the rating method. Regarding the experimental approach, 80 of the 186 reviewed studies used experimental observations in 35 different

primate species; 36% systematically assessed and found temporal stability (Carter et al., 2014; Fairbanks, 2001; Kutsukake et al., 2012; Tomassetti et al., 2019) and only 16% examined and found stability across situations (Dammhahn & Almeling, 2012; Massen et al., 2013; Uher, Addessi, et al., 2013). Finally, only 7% of the reviewed studies systematically examined and found temporal consistency using solely naturalistic observations and none of them assessed personality consistency across multiple distinct contexts.

In this chapter, the three different approaches that have been developed to collect data on nonhuman primate personality will be examined: rating method, experimental method, and naturalistic method. All three will be used in this doctoral research. Second, the different statistical methods used in the literature to estimate the reliability of the personality measurements will be presented. More specifically, we will focus particularly on the two most commonly used statistical methods (Intraclass Correlation Coefficient and Cronbach's alpha) and a less common statistical method (Generalizability Theory). The latter will be used throughout this research as this method allows to unify the different analyses required to test for temporal and contextual consistency of personality. Third, we will discuss the two different approaches (i.e., traditional and conceptual approaches) used in animal personality research to determine the personality structure of individuals – i.e., establishing the different traits that compose an individual's personality.

2.1 Assessment of nonhuman primate personality

Three methods have been developed to assess animal personality: ratings, experimental and naturalistic behavioural observations (King & Figueredo, 1997; Koski, 2011b; Kutsukake et al., 2012; Weiss, King, & Hopkins, 2007). In this section, I will

critically evaluate the different methods that have been used to measure animal personality, discussing the strengths and weaknesses of each.

2.1.1. Rating method

In nonhuman animal personality studies, two formats of items have been developed: adjective and behaviour-descriptive verb. First, the adjectives were initially developed for human personality research (Goldberg, 1990; McCrae & John, 1992), and were later adapted, so they could be used to determine different personality traits in various nonhuman species (Bard & Gardner, 1996; Clay et al., 2015; King & Landau, 2003; Stevenson-Hinde & Zunz, 1978; Uher & Asendorpf, 2008; Uher & Visalberghi, 2016; Weiss & King, 2015). Uher and colleagues (2008, 2013, 2016) developed the second format of items, namely behaviour-descriptive verb item, to measure different personality traits in captive great apes, capuchins and macaques. This type relies on the behavioural description of a specific personality trait occurring in a specific situation. Overall, the rating method requires familiar observers or carers to rate the individuals' personality using a series of adjectives or behavioural descriptors. The raters are asked to make their judgements based on their overall idea of how the individual of interest behaves.

By using the rating method that derived from human personality questionnaire and mostly from the Big-Five model (Goldberg, 1990; McCrae & John, 1992), psychological researchers studying animal personality, especially nonhuman primates, can, therefore, make interesting phylogenetic comparisons. More specifically, such comparisons can lead to a better understanding of the evolutive origins of human personality and a possible identification of the different selective pressures that may have occurred across evolution (Buss, 1988; Nettle, 2006).

Strengths

The adjective items have been reported to reflect daily behaviours. The behavioural actions expressed by the animals in their day-to-day life were found to match the personality profile obtained from the raters. For instance, dominance was positively correlated with agonistic behaviours (e.g., aggressive display) in captive chimpanzees (Pederson et al., 2005). This outcome validates, therefore, the reliability of the adjective items to assess animal personality (Capitanio, 1999; Freeman et al., 2013; Uher & Asendorpf, 2008).

When compared to personality traits generated from experimental behavioural observations, the traits generated from the behaviour-descriptive verb items showed more consistency than the ones that were created from the adjective items (Uher & Asendorpf, 2008). The behaviour ratings provide a detailed measure of the personality trait targeted as the items describe behaviours occurring in a specific situation. For instance, to assess curiosity, the authors used the following item: "Confronted with novel food, *Name* (usually) ignores it" (Uher & Asendorpf, 2008). The authors suggested that the adjective ratings might have a broader bandwidth and may be less informative than the behaviour ratings.

The rating method can be quick to administer in both captive (Baker et al., 2015; Weiss & King, 2015; Weiss et al., 2007) and wild (Ebenau et al., 2019; Tkaczynski et al., 2018; Weiss et al., 2017) settings allowing researchers to obtain large sample size, provided that the raters know the animals. With such a big sample, researchers can show a better representation of personality traits across the subjects.

The adjectives used to assess animal personality are often conceptually related to the personality trait targeted; some of the items used in the questionnaires often represent a clear reflection of what they are meant to measure. For instance, in the Hominoid Personality Questionnaire (King & Figueredo, 1997), "curious" or "innovative", which are undoubtedly used to measure curiosity, have been reported to positively load onto the component "openness" in various primate species (Eckardt et al., 2015; Morton et al., 2013; Weiss, Adams, Widdig, & Gerald, 2011). With such a conceptual characteristic, this type of adjective items is likely to cover a broad view of a personality trait, and may then cover multiple facets of the same trait.

Weaknesses

The rating approach garnered criticisms in personality research because it was claimed to be anthropomorphic and subjective (Uher, 2008a). People remember more major life events than day-to-day events (Koppel, Brown, Stone, Coman, & Hirst, 2013), and particularly negative over positive events (Kensinger & Schacter, 2006), and tend to turn to others for social support (Harlow & Cantor, 1995). Hence, it is possible that raters are influenced in their ratings if they witnessed a major event that recently occurred in the subject's life (e.g., big fight) or if they were told by other raters who witnessed this particular event, leading to either an underestimation or overestimation of the personality trait scores of an individual.

Additionally, the fact that the adjectives used in the animal questionnaires were derived from the human literature (Goldberg, 1990), some of them might have been less meaningful for the studied species. For instance, some questionnaires used adjectives such as "jealous" or "cool" to define the personality of primate species (Adams et al., 2015; Eckardt et al., 2015; Inoue-Murayama, Yokoyama, Yamanashi, & Weiss, 2018). Such

adjectives are probably more relevant to humans than animals; for instance, jealousy is measurable in the everyday life of human behaviours (Mullen, 1996). It is more difficult to extend such a view in animals; what does being jealous mean for a chimpanzee for instance? Having a short definition certainly helps the rater determine how this adjective could apply to the individual, but to what extent? What daily behaviours can the raters relate to when estimating how jealous or cool an individual is? Researchers assessing animal personality should perhaps be more careful with the use of anthropomorphic words or the selection of items when studying animals.

A last weakness of the rating method is perhaps the behavioural descriptor item. This type of items provides a descriptive assessment of the personality trait targeted as it is associated to a specific situation (Uher, 2008a, 2008b, 2011). However, if only a few questions are used to describe the personality trait of interest, this type of items is likely to cover only a definite facet of the personality trait.

2.1.2. Experimental method

Unlike the rating method, behavioural observations in an experimental setting reflects a more biological approach where the personality assessment relies only on the coding of behaviours displayed by the subjects (Freeman & Gosling, 2010; Koski, 2011a). Such a biological approach is objective and provides a more unbiased view and less chance of misinterpretation of inter-individual differences as long as experimental designs and clear behavioural definitions are used (Carter et al., 2013; Gosling, 2001). Using such a method of assessment requires the researcher to determine in advance the personality trait of interest. Indeed,. past research developed experimental assays to target specific personality traits, providing standardized environments in which to observe the different traits (Réale et al., 2007). For instance, in nonhuman primates, novel object tests are used to measure

exploratiVeness-related behaviours (Tkaczynski et al., 2018; Uher et al., 2008) whereas snake tests are used to measure boldness-related behaviours (Massen et al., 2013; Šlipogor et al., 2016). These well-designed experimental protocols can be applied to different social groups and species.

Strengths

One of the key advantages of using experiments is to have a controlled setting where one can manipulate a single variable to examine the responses from the subjects (Freeman & Gosling, 2010). By doing so, we can test specific hypotheses about animal personality and provide very detailed information about the personality trait targeted as the trait is measured in a specific experimental situation. Such manipulations of the experimental setting are likely to elicit specific types of behavioural reactions towards a stimulus or an environment across subjects (Freeman & Gosling, 2010; Freeman et al., 2011). For instance, novel object or novel food tests are used to examine various degrees of exploratiVeness-related behaviours such as gazing, approaching, touching, handling or moving away from the stimulus (Gosling, 2001; Massen et al., 2013; Réale et al., 2007). These types of experimental tests can be carried out in both solitary and social settings, depending on the research question one is investigating (Koski & Burkart, 2015; Massen et al., 2013; Šlipogor et al., 2016; Uher et al., 2008).

Like with the rating method, personality assessment using experimental observations is relatively quick to obtain as the animals are often observed in the same situation twice to measure temporal consistency (Dammhahn & Almeling, 2012; Fairbanks, 2001; James et al., 2007; Šlipogor et al., 2016; Tomassetti et al., 2019; Uher et al., 2008).

Weaknesses

A notable limitation of this approach is its level of ecological validity. The subjects are often tested in an artificial environment that does not always reflect real-life situations, such as the use of open field test for wild animals (Dammhahn, 2012) or novel objects for captive or wild animals (Tkaczynski et al., 2018; Uher et al., 2008). Some controlled experiments are designed to reduce all distractions or variations coming from the environment which may result in a lack of familiarity for the animals and influence their behavioural responses.

Additionally, using an experimental test that was originally developed to measure a personality trait in a specific species might sometimes lead to inaccurate conclusions when applied to another species. The same experiment might eventually target a different personality trait than the one originally planned (Carter et al., 2012b). In Carter and colleagues' study (2012b), based on previous studies, the authors used two experimental assays (i.e., novel food test and threatening stimuli test) to measure boldness in wild chacma baboons (*Papio ursinus*). Contrary to what the authors expected, the two tests did not measure the same trait; instead, the threatening stimuli test measured anxiety. The literature refers to this inaccuracy as the jingle-jangle fallacies, where two labels might be casually used to describe the same personality trait (jangle) or one single label is used to describe two distinct personality traits measured with different experimental tests (jingle) (Carter et al., 2013, 2012b). This is why it is important to define explicitly the personality traits based on the behaviours displayed by the subjects, but it is also important to use multiple tests to target the same trait to ensure that the right trait is targeted.

Furthermore, as a result of the experimental design that induces specific responses from the subject, research applying this method focused thus far on a limited number of

behaviours displayed in a single or very similar situations (e.g., novel food test to measure curiosity in novelty-like situations: Uher et al., 2008). Therefore, other spontaneous behaviours that could possibly emphasise subtle inter-individual variations are overlooked. Such a method is likely to target only a specific facet of a personality trait, as the experimental observations are closely associated with a specific situation.

One last limitation with this method is that experimental studies often require the researchers to capture, handle or mark the subjects. Such a procedure is likely to affect the behavioural responses of the subjects who are likely to show discomfort and distress during the experimental test (Putman, 1995).

2.1.3. Naturalistic method

As for the experimental method, behavioural observations in a natural setting reflects a more biological approach than the rating method (Gosling, 2008), as it also considers behavioural observations. However, unlike the experimental observations, naturalistic behavioural observations represent an objective method of high ecological validity as the personality assessment relies only on the coding of behaviours displayed by the subjects in day-to-day contexts, where no experimental manipulation has been carried out (Freeman & Gosling, 2011). Clear definitions of behaviours are also required beforehand in order to avoid any misinterpretations of personality assessment. By doing so, naturalistic observations are, therefore, likely to capture a broad overview of the individual's personality by targeting multiple behaviours expressed in various situations. More specifically, the individuals can interact with both their physical and social environments without any constraints imposed by the measuring method (Miller, 1977), providing, therefore, genuine behavioural responses of the individual.

Strengths

Contrary to some preconceived ideas that animals are only observed in one context (Freeman et al., 2011), observing the behaviour in the natural course of individuals' lives is likely to cover a wide range of distinct contexts (e.g., play, vigilance or aggression) and multiple behaviours over different time periods (Miller, 1977). It contributes to a more detailed description of an individual's personality traits in relation to the contexts they occur in (Mehta & Gosling, 2008; Uher, 2008a) and provides a better understanding of the complexity of personality expression as it seems to cover different facets of the same personality trait and possibly reveal inter-individual differences.

Like the rating and experimental methods, this approach of high ecological validity is beneficial for comparative studies. It can be applied to different social groups or species, and different type of settings (captive, wild), provided that similar behaviours and contexts are observable (e.g., feeding, resting, play) when comparing different species.

Weaknesses

One of the limitations of this method is the measurement of behaviours in natural settings. Unlike the experimental studies, the researcher cannot control any distractions or variations coming from the environment. Consequently, the animals may show fluctuations in their behavioural responses independently of their personality traits.

Of course, another drawback of this method is the way to collect the behavioural data: live coding versus recording. Regarding the live coding, a detailed coding scheme is not possible to use, and one is likely to miss behaviours during the observation period. With regard to recordings, although a detailed coding scheme is possible to use during the behavioural coding phase, it can be time-consuming to go through all recordings to code

the different behaviours of interest (Martin & Bateson, 2007). In both ways, the researcher does not have any control while collecting the observations. Indeed, if the focal subject decides to leave the visible area during the observation period, then the data are incomplete and likely unusable.

In comparison with the other two methods, the naturalistic method can be demanding and time-consuming depending on the level of details the researcher is looking for. The focal individual has to be followed for a relatively large amount of time on multiple occasions across different contexts allowing to capture a full range of behavioural variation across the subjects. A large dataset is often required to gather enough data allowing to test specific hypotheses. However, it is important to note that subjects with a small number of observations should not be removed from any analyses as it may reduce the chance to detect variance across the individuals (Martin, Nussey, Wilson, & Réale, 2011).

2.1.4. Relevance of the three methods for this thesis

Personality can be expressed across multiple situations or contexts using various behaviours over time (Allport, 1961; Gosling, 2001; Sih, Bell, Johnson, et al., 2004). To provide a comprehensive description of the personality profile of individuals, the use of multiple methods of assessment is likely to help capture the diversity of expression of different personality traits. Each of these methods presents different strengths and weaknesses, they complement one another in the sense that the flaws of one method can be addressed by the others. Using perhaps one method in combination with another could show us an improved understanding of personality.

Each method may provide an important contribution to the assessment of the personality trait targeted. Because the experimental approach obtains a limited number of

behaviours displayed in a specific situation at a specific time, this method is likely to target a specific facet of a personality trait. In contrast, the naturalistic approach captures multiple behaviours displayed across multiple daily contexts over time; thus, this method is likely to provide the broadest description of a personality trait. Finally, depending on the format of the rating, this approach is likely to provide both a broad (adjective format) and a detailed (behavioural descriptor format) description of a personality trait.

2.2 Reliability and validity of the assessment

One crucial step to take in research is assessing the reliability of the measurements to ensure the validity of the study (Martin & Bateson, 2007). Reliability refers to the degree to which measurement is repeatable and consistent, and validity implies the extent to which a measurement estimates what is intended to be measured (Martin & Bateson, 2007). When conducting reliability, it is important either to be blind to the aims of the study while collecting or coding the data or ask an independent person to code a sub-set of the data (Burghardt et al., 2012). Yet, this information is only briefly mentioned, if not ignored, in empirical studies (Burghardt et al., 2012; Kaufman & Rosenthal, 2009).

In animal personality research, different statistical methods have been used to test for temporal and contextual consistency, but two of them primarily dominate the area: Intraclass correlation coefficient (ICC) and Cronbach's alpha. Of course, one simple reason that could explain the application of these statistical tests is the use of continuous data when examining temporal or contextual repeatability in behavioural studies or assessing inter-rater reliability in rating studies. Other methods have also been used, such as Pearson correlation, Spearman rank correlation, or Kendall coefficient (Dutton, 2008; Seyfarth et al., 2012; Uher et al., 2008). All of these statistical tests are often referred to as Classical Test Theory (Brennan, 2011) which partitions observed-score variance into two

types: true-score variance and undifferenciated error variance (Shavelson, Webb, & Rowley, 1989). The issue garnered with this approach is that these tests do not distinguish between the different sources of error that may occur when collecting data (e.g., repetition of measurements, sampling measurement). In contrast, generalizability theory has been developed to address these concerns enabling, therefore, the identification of individual sources of error and their interactions that may emerge (Brennan, 2011; Shavelson & Webb, 1991; Shavelson et al., 1989). In this section, I will discuss three main methods – intraclass correlation coefficient, Cronbach's alpha and generalizability theory – that are used to assess the repeatability of the measurements in animal personality.

2.2.1. Intraclass correlation coefficient

ICC measures the proportion of total variance in behaviours or items that is due to differences between subjects while considering the within-subject variance (McGraw & Wong, 1996; Shrout & Fleiss, 1979). Shrout and Fleiss (1979) reported three types of ICC: 1) Each target is rated by a different set of k raters, randomly selected from a larger population of raters – ICC(1) or one-way random; 2) A random sample of *k* raters is selected from a larger population, and each rater rates each target, that is, each rater rates ***n*** targets altogether – ICC(2) or two-way random; 3) Each target is rated by each of the same *k* raters, who are the only raters of interest – ICC(3) or two-way mixed.

In behavioural studies, the ICC(3) is often used to assess the temporal consistency of the variables (Dammhahn, 2012; Koski, 2011b; Massen et al., 2013), though it can be also occasionally used to measure the contextual consistency (Šlipogor et al., 2016); the same procedure is applied whether we are interested in measuring temporal or contextual consistency. A two-way mixed model – ICC(3,1) – is used with the time period as the fixed variable while the individual is the random variable. Usually, a behaviour is

considered consistent over time if the ICC value is significantly different from 0 ($p < 0.05$). It is important to note that some researchers may instead ignore the p-value which may be perceived as too strict (Weiss, 2017).

In rating studies, the inter-rater reliability is measured using two intraclass correlation coefficient: ICC(3,1) and ICC(3,*k*). ICC(3,1) is used to estimate the reliability of ratings of any single rater which establishes how accurate a single rater would be if the ratings were only performed by them. This type is to be considered when comparing the values between different studies or groups, in case the number of raters differs. ICC(3,*k*) estimates the reliability of the mean ratings across *k* raters (Shrout & Fleiss, 1979), and is used when deciding which items to retain for further analyses. Some studies use a strict cut-off to decide whether an item should be retained (Baker et al., 2015; Tkaczynski et al., 2018) whereas most of the studies are more flexible and retain any items that have an ICC value above 0 (King, Weiss, & Farmer, 2005; Konečná et al., 2012; Weiss et al., 2015). The latter approach assumes that lower inter-rater reliabilities may indicate the rarity of items, thus they should not be ignored; if the items are instead a random error, then they should not load on any components in the following analyses (Weiss, 2017).

2.2.2. Cronbach's alpha

Cronbach's alpha allows to assess internal consistency, which describes the extent to which different contexts (for behavioural studies) or items (for rating studies) measured the same underlying behaviour or personality trait (Bland & Altman, 1997; Cronbach, 1951). The Cronbach's alpha is obtained from the number of contexts/items that are included in the study, the average inter-context/item covariance among the contexts/items and the average variance. The value increases if the number of contexts/items or the average inter-item correlation increases.

As frequently used in psychology when developing a scale (Peterson, 1994), any value equal or above 0.7 would suggest that the same behaviour or personality trait is reliable across the contexts. If the alpha value is below, a stepwise approach is used to re-estimate the value Cronbach's alpha if a particular context is removed. Consequently, if one of the values after deletion is greater than the original value, then the context is removed, and the analysis is re-run. This process is usually repeated until there are no alpha values after deletion greater than the overall alpha (Field, 2013).

2.2.3. Generalizability theory

Generalizability Theory (GT) is a statistical theory for evaluating the dependability of behavioural measurements (Brennan, 1992, 2011; Shavelson & Webb, 1991; Shavelson et al., 1989). In this thesis, GT is used to assess the extent to which the behaviour or personality trait scores generalize across both contexts and time. GT uses ANOVA methods to partition the score of an individual and disentangle multiple sources of error that contribute to the undifferentiated error in classical reliability tests (Brennan, 2001; Shavelson & Webb, 1991). More specifically, although this statistical theory is not widely used in animal personality studies, unlike the ICC or Cronbach's alpha, GT allows to identify multiple sources of systematic and unsystematic errors and estimates the variance components associated with each source of variation in a single analysis (Shavelson & Webb, 1991). For instance, in a random crossed design – subject by context by time – the subject represents the object of measurement which is not a source of error but a systematic source of variance. Both context and time represent the facets of generalization, which are potential sources of measurement error (see Figure 2.1 and Table 2.1). By using GT, a researcher is able to establish how many contexts and time periods are needed to acquire dependable scores.

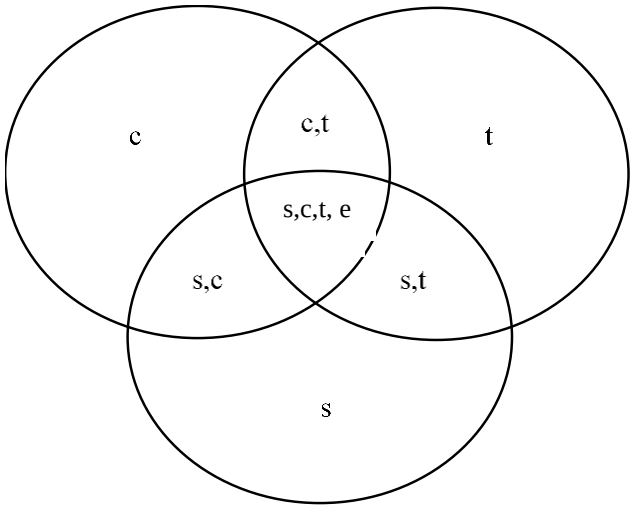


Figure 2.1. Venn diagram representing the different sources of variance for a random crossed design, subject (s) by context (c) by time (t).

Table 2.1. Overview of the different sources of variance in the random crossed design: subject (s) by context (c) by time (t).

Source of variance	Explanation	Variance component
s	Universe of admissible observations (true score) for subject averaged over context and time; a high variance component indicates that the subjects differ from one another	σ^2_s
c	Effect for context averaged over subject and time; a high variance component indicates that the expression of behaviour or personality trait differs between the contexts	σ^2_c
t	Effect for time averaged over subject and context; a high variance component indicates that the expression of a behaviour or personality trait differs between the time periods	σ^2_t
s x c	Idiosyncratic response of subject by context, averaged over time; a high variance component indicates that there are inter-individual differences in the contextual consistency of a behaviour or personality trait	σ^2_{sc}
s x t	Idiosyncratic response of subject by time averaged over context; a high variance component indicates that there are inter-individual differences in the temporal consistency of a behaviour or personality trait	σ^2_{st}
c x t	Idiosyncratic response of context by time averaged over the subject; a high variance component indicates that there are inter-individual differences in both contextual and temporal consistency of a behaviour or personality trait	σ^2_{ct}
s x c x t, e	Residual effect involving the triple interaction and all other sources of error not explicitly represented in the universe of admissible observations	$\sigma^2_{sct,\ e}$

When used in animal personality research, GT can bring interesting and detailed information regarding the consistency aspect of the measured behaviours or personality traits. Indeed, unlike CTT (i.e., ICC, Cronbach's α), GT provides the partitioning of the observed scores into the different effects (i.e., subject, context and time) as well as their interactions with one another [i.e., context by time – constant effect for all subjects due to differences in context characteristic from one time period to another; Shavelson & Webb, 1991)] and with the subjects (i.e., subject by context or subject by time). The subject interactions can tell the researcher whether some individuals show more consistency than

others across contexts and/or time, information that cannot be obtained with classical reliability tests (Brennan, 2001; Shavelson & Webb, 1991).

In GT, two elements are included, generalizability (G) study which is an estimation of all the sources of variance that influence the measurements, and decision (D) study which gives the optimal design to obtain high reliability. D-study uses the information from G-study to test different designs in order to determine the most optimal design that would maximise the reliability of the measurements (i.e., minimising the measurement error). For instance, D-study aims to answer questions such as "Would an increase in the number of contexts or time periods improves the reliability of the design?". The variance estimate for each effect is calculated using the mean squares (see Table 2.2), and the percentage of the total variance for each variance estimate is computed by dividing each variance estimate by the total variance (Brennan, 2001). The estimation of variance components can be subjected to some variability in the sampling method. Consequently, some variance components may sometimes take on negative values, although variance components are nonnegative by definition (Brennan, 2001; Shavelson & Webb, 1991). Such negative variances may emerge as a result of misspecification of the model leading often to large negative variance component or sampling error (i.e., small sample) leading often to smaller negative variance component (i.e., close to 0). Consequently, large negative variance component would require the researcher to respecify the measurement model by dropping some of the facets (Shavelson & Webb, 1991). Conversely, in case of smaller negative values, an approach commonly accepted in research is to replace the negative estimate with 0 to prevent the use of negative values (conceptually impossible) in the calculation of the different effects (Brennan, 2001; Shavelson & Webb, 1991).

Table 2.2. Overview of the estimators of variance components for the G-study subject (s) by context (c) by time (t) design and subject (s) by time (t) design.

Effect (s x c x t design)	Formula
s	$\frac{MS(s) - MS(sc) - MS(st) + MS(sct)}{n_c n_t}$
c	$\frac{MS(c) - MS(sc) - MS(ct) + MS(sct)}{n_s n_t}$
t	$\frac{MS(t) - MS(st) - MS(ct) + MS(sct)}{n_s n_c}$
s x c	$\frac{MS(sc) - MS(sct)}{n_t}$
s x t	$\frac{MS(st) - MS(sct)}{n_c}$
c x t	$\frac{MS(ct) - MS(sct)}{n_s}$
s x c x t, e	$MS(sct)$
Effect (s x t design)	**Formula**
s	$\frac{MS(s) - MS(st)}{n_t}$
t	$\frac{MS(t) - MS(st)}{n_s}$
s x t, e	$MS(st)$

MS denotes the mean square; n_s denotes the number of individuals, n_c and n_t denote the number of levels of the facet context and time, respectively.

Additionally, GT provides two reliability-like coefficients: generalizability (G) and dependability (ϕ) coefficients (Brennan, 2003). The former is the analogue of a reliability coefficient in classical test theory; the G-coefficient represents how the behaviour or personality score of an individual relates to other individuals (between-individual reliability). It is defined as

$$G_{s\ x\ c\ x\ t\ design} = \frac{\sigma^2(s)}{\sigma^2(s) + \frac{\sigma^2(sc)}{n_c} + \frac{\sigma^2(st)}{n_t} + \frac{\sigma^2(sct)}{n_c n_t}} \qquad G_{s\ x\ t\ design} = \frac{\sigma^2(s)}{\sigma^2(s) + \frac{\sigma^2(st)}{n_t}}$$

The latter represents how the behaviour or personality score of an individual is regardless of the other individuals' behaviour or personality score (within-individual reliability). It is defined as

$$\phi_{s\ x\ c\ x\ t\ design} = \frac{\sigma^2(s)}{\sigma^2(s) + \frac{\sigma^2(sc)}{n_c} + \frac{\sigma^2(st)}{n_t} + \frac{\sigma^2(sct)}{n_c n_t} + \frac{\sigma^2(c)}{n_c} + \frac{\sigma^2(t)}{n_t} + \frac{\sigma^2(ct)}{n_c n_t}}$$

$$\phi_{s\ x\ t\ design} = \frac{\sigma^2(s)}{\sigma^2(s) + \frac{\sigma^2(st)}{n_t} + \frac{\sigma^2(t)}{n_c n_t}}$$

For the purposes of the current research, only the G-coefficient will be reported throughout the empirical chapters, as I am interested in how reliable the behaviour/personality trait is overall, that is, how an individual's score is reliable in comparison with the other individuals' scores.

2.3 Determining the personality structure

The personality structure of individuals (i.e., different personality traits composing the personality of individuals) can be determined using two different approaches: traditional and conceptual approaches. The two include a different perspective when describing the structure. The traditional approach works on the behavioural level where only reliable behaviours/items are considered for the analysis. The conceptual approach works on the trait level where all behaviours that were formerly deemed relevant are included. In this section, I will discuss how these two approaches operate to determine the personality structure of individuals.

2.3.1. Traditional approach

Both human (Goldberg, 1990; John & Srivastava, 1999) and nonhuman animal (Bergvall, Schäpers, Kjellander, & Weiss, 2011; Budaev, Zworykin, & Mochek, 1999; Massen et al., 2013) studies tend to use factor analysis to determine the personality structure. This statistical tool allows researchers to uncover an underlying personality trait by often reducing a larger set of variables (behaviour, item) into a set that shares a common variance. After a series of correlations between the variables and transformations, a new set of linear combinations of the original variables is created (i.e., components or factors).

All original variables that share the same variance load onto the same component/factor; the loading reflects the correlation between the original variables and the component factor (Budaev, 2010). There are two types of methods that can be used to determine the personality structure: principal component analysis (PCA) and factor analysis (FA); both methods are very similar in terms of achievement. On one hand, PCA is a dimensionality reduction method which is most appropriate when the main objective is to reduce the number of dimensions. On the other hand, FA aims to measure unobservable latent constructs (i.e., conceptual characteristic of an individual's personality profile) that explain the correlation between the variables (Budaev, 2010).

Different criteria must be met before considering the solution obtained as stable and reliable. First, one must check the Kaiser-Mayer-Olkin (KMO) measure of sampling adequacy, which represents the proportion of variance in the variables that might be caused by underlying factors. This statistic is calculated for each variable (e.g., behaviour, item) and for the complete set of variables. Second, Bartlett's test of sphericity tests whether the variables are related or not. Both measures indicate whether the dataset is robust enough to be subjected to a factor analysis and is likely to provide a reliable solution. In animal

behavioural studies, a KMO above 0.5 is strongly recommended (Budaev, 2010) and Bartlett's test has to be significantly different from 0 indicating that the variables are correlated highly enough to provide a strong basis to carry out the factorial analysis (Loewen & Gonulal, 2015). If these two measures do not meet the criteria, then the researcher may consider removing some variables based on the individual KMO or relevance of the variable included in the study.

Then, a factor rotation must be performed on the data in order to optimize the loading of the original variables onto the different components/factors. There are two types of rotation: orthogonal and oblique. The first type does not allow the components/factors to correlate, unlike the second type. Often, both rotations are applied and the determining of which one of the two is the most appropriate is based on the correlation values between the different components/factors (oblique rotation). If the values are low, then the orthogonal rotation is retained for further interpretation. The number of components/factors to retain can rely on the eigenvalue (>1) and scree plot (Budaev, 2010). Occasionally, lower eigenvalues can be considered if the components extracted explain at least 80% of the variance (Field, 2013). After finding the right solution, the researcher has to label the different components/factors based on the variables' loadings as well as the socioecology of the species studied and can also rely on the personality structure of a closely related species (Weiss, 2017). This empirical approach has been mostly used in the nonhuman primate literature to determine the personality structure of the species studied, regardless of the method of assessment used (ratings, experimental or naturalistic observations) (Konečná et al., 2012; Koski, 2011b; Massen et al., 2013; Seyfarth et al., 2012; Šlipogor et al., 2016; Weiss & King, 2015). In this thesis, the traditional approach was used to analyse the data where only consistent behaviours were included in the analyses allowing the identification of the personality structure.

2.3.2. Conceptual approach

To determine the different personality traits targeted, Uher (2008a,b) developed a different approach, which is more conceptual compared to the traditional approach described above. The conceptual approach was developed for the rating and observational methods; the selection of the items/behaviours to reflect the personality traits of interest relied on a thorough review of the literature (Uher, 2008a, 2008b, 2011b). This approach was applied to different nonhuman primate species (Uher, 2008a, 2008b, 2011a; Uher & Asendorpf, 2008; Uher et al., 2008; Uher & Visalberghi, 2016; Uher, Werner, et al., 2013).

The personality traits generated with this approach are derived from observable and measurable behaviours occurring in specific situations (Uher, 2008b, 2008a, 2011a, 2011b). This behavioural repertoire approach is to some degree a detailed reflection of the personality traits of the species studied as it relies on both universal behaviours (e.g., play action, self-scratch) and species-specific behaviours (e.g., grooming in chimpanzees, coercive hand-holding in orangutans). All selected behaviours/items have been predetermined to belong to certain personality traits. Consequently, each of the behaviours/items that reflects the same personality trait is then merged to generate a trait construct, which can be tested for temporal and contextual consistency (Tomassetti et al., 2019; Uher et al., 2008; Uher, Werner, et al., 2013).

Based on scientific observations and analyses (e.g., reliability, factor analysis), any scientifically described behaviour occurring in different situations is, therefore, selected and represents an indicator of the targeted personality trait (Uher, Addessi, et al., 2013). Therefore, this approach is more inclusive in comparison with the more traditional approach that includes only reliable behaviours/items.

3. Personality consistency measured using naturalistic observations across daily contexts in sanctuary chimpanzees (Pan troglodytes)

Abstract

Personality plays a central role in individuals' daily life by influencing their decision making or relationships with others. Personality is defined as inter-individual differences in behaviour-based traits that reflect components of the self. Chimpanzees' personality has been mostly studied using rating or experimental methods. Using naturalistic observations to measure personality in daily contexts may provide a better understanding of individuals; they can interact with both physical and social environments without any constraints in terms of their behavioural expression. We examined the behavioural consistency of 22 sanctuary chimpanzees across multiple distinct contexts and over two time periods (4 years apart). A coding scheme was developed which included 22 behaviours and eight ecologically relevant contexts. Overall, six to seven behaviours were found to be consistent across two to four contexts and over time. The principal component analysis revealed four components: exploratiVeness, anxiety, boldness and sociability. The approach used in this study focuses on different trait-relevant behaviours and ecologically relevant contexts for the species, allowing us to highlight inter-individual differences in multiple distinct contexts. Future research would benefit from a combination of such an approach with rating and/or experimental methods, as they may provide a finer description of the personality structure of nonhuman primate species. Considering a wide range of behaviours and contexts, which are frequently expressed in various primate species, can

help better assess personality, highlight possible variations across species and provide, thus, interesting insight into the evolutionary roots of human personality.

3.1 Introduction

Personality is defined as inter-individual differences in behaviour-based traits (Allport, 1961; Allport & Odbert, 1936; Funder, 2001). In humans, individuals differ in the way they express their personality traits; their behaviours may fluctuate when interacting with both their physical and social environments in the natural course of their lives (Ickes, Snyder, & Garcia, 1997; Mehl, Gosling, & Pennebaker, 2006). An individual's personality traits, in general, are expected to show some consistency over time (Roberts & DelVecchio, 2000; Weisbuch, Slepian, Clarke, Ambady, & Veenstra-VanderWeele, 2010) as well as across different situations (Funder & Colvin, 1991; Furr & Funder, 2004; Moskowitz, 1982; Sherman, Nave, & Funder, 2010; Weisbuch et al., 2010). Measuring personality traits that can be expressed in the daily life of nonhuman primates using behavioural observations could help identify variations, and possibly provide an important contribution to better understand the evolutionary history of human personality (Buss, 1988; Nettle, 2006).

As in humans, research on nonhuman primate personality also considered both contextual and temporal consistency to be key criteria to examine when assessing personality (Brent et al., 2014; Massen et al., 2013; Weiss et al., 2007). Consistency has been established using rating methods (King & Figueredo, 1997; King et al., 2008; Weiss & King, 2015), experiments (Koski & Burkart, 2015; Kutsukake et al., 2012; Massen et al., 2013) and only to a small extent via naturalistic observations (Koski, 2011b; Seyfarth et

al., 2012; Tkaczynski et al., 2018). Temporal consistency was found for all three personality methodologies (see Chapter 2; Freeman & Gosling, 2010; Gosling, 2001); naturalistic observations revealed temporal consistency for personality traits (e.g., sociability, boldness, anxiety, boldness) in tufted capuchin (*Sapajus paella:*(Byrne & Suomi, 1995), four macaque species (*Macaca mulatta*: Bardi, Shimizu, Fujita, Borgognini-Tarli, & Huffman, 2001; Brent et al., 2014; Maestripieri, 2000; von Borell et al., 2016); *M. nernestrina:* Reite & Short, 1980*; M. nigra*: Neumann, Agil, Widdig, & Engelhardt, 2013; *M. sylvanus*: Bardi et al., 2001; Tkaczynski et al., 2018), chacma baboons (*Papio ursinus*: Seyfarth et al., 2012) and in chimpanzees (Koski, 2011b). In contrast, contextual consistency was found with the experimental method for a range of personality traits across nonhuman primate species (Dammhahn & Almeling, 2012; Šlipogor et al., 2016; Uher, Addessi, & Visalberghi, 2013; Uher et al., 2008), such as sociability, boldness, exploratiVeness and anxiety in chimpanzees (Kutsukake et al., 2012; Massen et al., 2013; Uher et al., 2008). To our knowledge, contextual consistency has not yet been examined using a naturalistic method.

So far, naturalistic methods have been primarily used to validate or complement rating scales or supplement experimental measures in the assessment of personality consistency (Clay, Bloomsmith, Bard, Maple, & Marr, 2015; Freeman et al., 2013; Pederson, King, & Landau, 2005). It is important to note that these different methods to measure personality are likely to provide different information about a personality trait, and target, perhaps, different facets of the same personality trait (Carter et al., 2012a). However, a naturalistic approach offers important benefits besides showing high ecological validity. Specifically, this approach may include multiple distinct contexts (e.g., play, vigilance, feeding) and such a range of contexts is likely to contribute to a detailed

description of a personality trait (Gosling, 2008; Mehta & Gosling, 2008; Uher, 2008a). On the contrary, an experimental approach (Réale et al., 2007) or the use of behaviour-descriptive items in the rating approach (Uher, 2011b; Uher & Asendorpf, 2008), for instance, are tied to a specific situation, which may both target a specific aspect of a personality trait by limiting its measurement to few behaviours.

Naturalistic observations could contribute to a better understanding of personality traits in nonhuman primates in relation to the contexts they occur. There are few nonhuman primate studies that measured personality using naturalistic observations in various contexts (Brent et al., 2014; Konečná et al., 2008, 2012; Koski, 2011b; Neumann et al., 2013; Pederson et al., 2005; Schaefer & Steklis, 2014; Seyfarth et al., 2012), but none of them assessed personality consistency across multiple distinct contexts. One exception could be noted for Capitanio (1999)'s study – the author determined the personality of adult male rhesus macaques using the rating approach and correlated each personality dimensions with behaviours measured in different situations (experimental and observational).

Depending on the method of assessment (rating, experimental or naturalistic observations), a series of adjectives based on human research (Goldberg, 1990) or behaviours displayed by the subjects in an experimental or natural setting (see Chapter 2) were used to determine personality in nonhuman primates. All adjectives or behaviours sharing a common variance correlate and load on the same component (i.e., dimension) (Budaev, 2010). Past research described chimpanzees' personality by six dimensions - dominance, extraversion, conscientiousness, agreeableness, neuroticism, and openness (King & Figueredo, 1997; Pederson et al., 2005) – which, except for dominance, are referred to as the Big-Five in human personality research (Costa & McCrae, 1992). Other

studies investigated similar personality traits but used different labels (Réale et al., 2007) such as exploration tendency as an equivalent of openness, or sociability as an analogue name for extraversion and agreeableness (Koski, 2011b; Massen et al., 2013; Uher, Asendorpf, & Call, 2008).

Based on past personality and behavioural research in nonhuman primates, our study focused on behaviours that referred to some extent to four personality traits in chimpanzees. These behaviour-based traits have been both reliably rated (Clay et al., 2015; Freeman et al., 2013; King et al., 2008; Pederson et al., 2005; Weiss & King, 2015) and measured behaviourally by human observers (Anestis, 2005; Baker & Aureli, 1997; Koski, 2011b; Massen et al., 2013; Uher, Asendorpf, & Call, 2008): sociability, boldness, explorativeness and anxiety. Sociability is defined as closeness to others and can be observed for instance, in social proximity or grooming (Eckardt et al., 2015; Koski, 2011b; Pederson et al., 2005). Boldness is the willingness to engage in potentially harmful situations and can be observed through actions such as hitting, chasing or risky approaches (Anestis, 2005; Massen et al., 2013; Nishida et al., 1999). Explorativeness, defined as showing interest in objects or conspecifics, can be observable in gazing, combined with actions such as touching, approaching and manipulating (Forss et al., 2015; Kutsukake et al., 2012; Schuppli et al., 2017). Finally, anxiety, the fearful, stressful or tense responses towards potential dangers, is measurable using behaviours such as self-directed or vigilant behaviours (Baker & Aureli, 1997; Kutsukake et al., 2012; Uher et al., 2008). These traits are ecologically and evolutionary relevant (Réale et al., 2007; B. R. Smith & Blumstein, 2008) as they are expressed naturally in different primate species suggesting their universality (Uher, 2008a), and they are argued to affect their individual fitness (e.g., life span: Altschul et al., 2018; survival: Silk, Alberts, & Altmann, 2003).

The aim of the current study was to test whether the behaviours of 22 sanctuary chimpanzees at Chimfunshi Wildlife Orphanage (CWO), Zambia, are consistent across multiple distinct daily life contexts in natural settings. Based on previous experimental studies in captive chimpanzees (Kutsukake et al., 2012; Massen et al., 2013; Uher et al., 2008), we hypothesise that the behaviours are consistent across naturally occurring contexts. In addition, like previous research in chimpanzees (Koski, 2011b; Uher et al., 2008), we tested if there was a temporal consistency across two time periods (4 years apart). We developed a coding scheme based on previous behavioural studies in primates (e.g., Anestis, 2005; Kutsukake et al., 2012; Seyfarth et al., 2012; Šlipogor, Gunhold-de Oliveira, Tadić, Massen, & Bugnyar, 2016) to capture a full range of behaviours – which targeted to some extent four personality traits – and ecologically relevant contexts.

3.2 Methods

Subjects and study site

Twenty-two chimpanzees (8 females) living at Chimfunshi Wildlife Orphanage (Zambia) were included in this study, with an age range of 5 and 32 years (mean ± SD: 16 ± 9.62). The subjects were members of two stable, multi-male multi-female colonies, showing natural fission-fusion dynamics. The larger colony (Colony 1) comprised 25 chimpanzees in a 77-hectare enclosure, whereas the second colony (Colony 4) included 11 chimpanzees in a 25-hectare enclosure. Due to their involvement in another behavioural study, only 13 and 9 members of each colony, respectively, were included in this current study. Additionally, due to the death and birth of members in each colony, the sex and age ratios differed slightly between 2013 and 2017 (see Table B.1).

Each enclosure contained naturally developed fruit groves, grasslands and forests in the miombo woodland, as well as an indoor area (used for midday feeding or medical check-up). Approximately 1.5 kilometres separated the two colonies preventing the subjects to observe chimpanzees from the other colony. The chimpanzees were provisioned with food (e.g., local fruits and vegetables) inside around noon and outside in the afternoon each day, and they could also forage in the forest. Water was always available outside via a water fountain.

The formation of the colonies took 2 to 5 years and the last colony formation ended at least 11 years before data collection. The two colonies were composed of a mix of wild-born chimpanzees and chimpanzees born at CWO. The colonies were formed by arrival dates of the wild-born chimpanzees rather than their geographic background. The wild-born chimpanzees were brought to CWO, individually or in pairs, from countries where wild chimpanzees live (e.g. Tanzania, Uganda and Rwanda). If they were born in the countries they were sent from, then the subspecies representation for these individuals would be 42–65 % for *Pan troglodytes schweinfurthii* and 31–42 % for *P. t. troglodytes* across the colonies (Wilson et al. 2008; Tutin et al. 2008). Although we do not know with certainty whether these chimpanzees were born in these countries, the colonies are likely composed of a comparable mix of sub-species, with no apparent phylogenetic differences. Most of the chimpanzees born at CWO are mother-reared; few were temporarily removed for health check but were put back with their conspecifics as soon as all concerns were cleared up.

Data collection

Video-recordings were collected using focal animal sampling (Altmann, 1974); the subjects were followed for 4 min (2013) or 5 min (2017), once or twice a day during the morning (from 7:30 to 11:30 am) and/or afternoon (from 1 to 5:30 pm) sessions. Recording occurred over two field seasons, between June and September 2013 and between May and September 2017. Prior to each recording session, the order of focal animal sampling was randomised to avoid any bias towards the same individual and provide a balance between time periods and contexts. Thus, all subjects were observed at different time throughout the day and in different contexts (social and non-social). Each conspecific was individually identified by the observer as soon as they were present within 10 meters of the focal subject. A total of 53 hours of recording was collected for this study, with approximately 2.5 hours per individual (Mean ± SD = 2.43 ± 0.50 hours).

Behavioural coding

The video-recordings obtained for this study allowed us to develop a very detailed coding scheme focusing on two parts: behaviours and context.

From frame by frame coding of videotapes, 22 behaviours were measured (see Table 3.1) and were based on previous personality and behavioural studies in nonhuman primates which reflected to some extent four common personality traits: sociability (Ebenau et al., 2019; Eckardt et al., 2015; Freeman et al., 2013; Koski, 2011b; Neumann et al., 2013; Silk, Brosnan, Henrich, Lambeth, & Shapiro, 2013; Suomi et al., 1996; van Hooff, 1973), boldness (Clay et al., 2015; Dammhahn & Almeling, 2012; Freeman et al., 2013; Koski & Burkart, 2015; Nishida et al., 1999; Santillán-Doherty et al., 2010), exploratiVeness (Damerius, Graber, et al., 2017; Forss et al., 2015; Massen et al., 2013; Santillán-Doherty

et al., 2010; Schuppli et al., 2017; Uher et al., 2008) and anxiety (Aureli & Waal, 1997; Baker & Aureli, 1997; Clay et al., 2015; Kalin & Shelton, 2003; Kutsukake, 2003; Kutsukake et al., 2012; Uher et al., 2008). By using video-recordings, we were able to code the behavioural actions independently from the contexts. The same action was not considered as part of the same series if there was a break of at least 5 seconds. This approach allowed to ensure the independence of the occurrences and avoid a possible inflated estimation of inter-individual differences. For each type of behaviour, we computed frequency per hour of context. Each subject obtained one single score per behaviour for each context, and each score was standardized as a z-score.

Table 3.1. Behaviours and their definitions associated with the four personality traits of interest. The table also depicts an overview of previous studies on nonhuman primates that mentioned these different behaviours.

Personality trait	Behaviour	Definition	Studies
Sociability	Groom	The subject or conspecific looks through a conspecific or subject's hair while picking at dirt or others	(Ebenau et al., 2019; Eckardt et al., 2015; Freeman et al., 2013; Koski, 2011b; Neumann et al., 2013; Silk et al., 2013; Suomi et al., 1996; van Hooff, 1973)
	Playful contact	The subject or conspecific initiates a social interaction with a conspecific or subject	
	Food sharing	The subject or conspecific gives food to, takes food without any resistance from another conspecific or subject or tolerates a conspecific or subject to take food from mouth to mouth or hand to hand	
	Food begging	The subject or conspecific is requesting food using hands or approaching his/her face towards conspecific/subject	
	Body contact	The subject or conspecific touches gently the conspecific or subject's body, presents his/her arm when approaching, embraces a conspecific or subject, presents hand or finger in conspecific or subject's mouth, the subject or conspecific's mouth enters in contact with other's mouth or with other's body part	
Boldness	Rough action	The subject jumps on conspecific, usually on the back, hits a conspecific with hands/feet, bites a conspecific, pulls a conspecific towards him/her or pushes a conspecific away with either hands/feet, shakes an object in the direction of a conspecific, shakes a conspecific's limb or swings own limb	(Clay et al., 2015; Dammhahn & Almeling, 2012; Freeman et al., 2013; Santillán-Doherty et al., 2010)

		towards a conspecific or stomps the ground with hands/feet in direction of a conspecific	
	Chase	The subject follows a conspecific (walks in a direct manner or runs) while the conspecific moves away	
	Display	The subject shows a tense posture while moving around in a perturbed manner	
	Risky action	The subject grabs edible food within arms' reach from a bigger/same size conspecific who can see the subject or steals edible food from a bigger/same size conspecific who resists in some ways	
	Throw	The subject throws objects towards another conspecific	
	Risky approach	The subject approaches while gazing conspecific(s) which put him/her in social danger	
Explorativeness	Gaze	The subject clearly directs gaze towards an object (e.g., stick, leaf, rock, rope, inedible food) or conspecific(s) (who are involved in a social interaction, eating or manipulating an object) for at least 2 seconds. Part of the subject's face needs to be visible to be counted	(Damerius, Graber, et al., 2017; Forss et al., 2015; Massen et al., 2013; Santillán-Doherty et al., 2010; Schuppli et al., 2017; Uher et al., 2008)
	Gaze Approach	The subject moves towards an object or conspecific(s) while focusing on it. Part of the subject's face needs to be visible to be counted	

	Gaze Touch	The subject's hand or foot enters into contact with an object or conspecific(s) while focusing on it. Part of the subject's face needs to be visible to be counted	
	Gaze Manipulation	The subject manipulates an object or join/hold conspecifics (involved in social interactions) while focusing on it	
Anxiety	Yawning	The subject opens widely his/her mouth	(Aureli & Waal, 1997; Baker & Aureli, 1997; Clay et al., 2015; Kalin & Shelton, 2003; Kutsukake, 2003; Kutsukake et al., 2012; Uher et al., 2008)
	Self-scratch	The subject rakes his/her hair with fingernails including mainly fingers or hand' s movements (gentle scratch) or including large arm movements (rough-scratch)	
	Self-touch	The subject touches his/her face or body with hand	
	Escape	The subject leaves hurriedly	
	Vigilance	The subject approaches slowly conspecific, stops occasionally to look around, grabs some food hurriedly, keeps looking around and is jumpy or freezes with no movements or vocalizations and shows tense posture	
	Pilo-erection	The subject's hair rises up	
	Rocking	The subject moves back and forth while sitting	

In experimental studies, a situation is determined by the presence of an object or experimental device placed in the environment surrounding a subject. For instance, a snake can be hidden in the enclosure which, then, corresponds to a predator-like situation (Koski & Burkart, 2015; Massen et al., 2013; Šlipogor et al., 2016). Here, we used a similar approach but instead of the object, we considered the conspecific(s) as the key cue to define the context. More specifically, the context was coded frame by frame and defined according to the presence and activity of the conspecifics (not of the subject) present within 10 metres of the focal. To be considered, a context had to be displayed by at least half of the conspecifics surrounding the focal subject and last at least 10 seconds (from the first behavioural indicator defining the context displayed by the conspecifics). The duration requirement was relaxed for three of the contexts (play, aggression, and vigilance; see table 3.2) which allowed us to test for contextual consistency of the personality-linked behaviours of the chimpanzees.

We classified the observations into eight naturally occurring and highly diverse contexts: feeding, play, grooming, resting, solitude, vigilance, aggression, and locomotion contexts (see Table 3.2). Some contexts rarely occurred in our data, such as object manipulation or copulation, and were not assessed further. Furthermore, only the first five contexts (feeding, play, grooming, resting, and solitude) were retained for the analyses because the last three were rarely observed in the data and represented less than 10% of the total duration of the complete dataset (see Table B.2). As previously done in personality studies (Kuhar, Stoinski, Lukas, & Maple, 2006; Pederson et al., 2005), grooming and play were merged to create an affiliative context as they together represented approximately 10% of the overall dataset (see Table B.2 for details). We conducted additional statistical

analyses when considering only the two most predominant contexts (feeding and resting) which represented more than 25% of the overall dataset in terms of duration.

Table 3.2. Definition of the analysed contexts.

Context	Definition
Feeding	The conspecifics are eating food provided by the sanctuary or are foraging in the enclosure
Play	The conspecifics are involved in a playful interaction
Grooming	The conspecifics are looking through others' hair
Resting	The conspecifics are sitting, lying down and/or self-grooming
Solitude	No conspecifics are present within 10 metres proximity
Vigilance	The conspecifics are alert while paying attention to events (including food being prepared by keepers) happening inside or outside the enclosure (e.g., vocalisations in background)
Aggression	The conspecifics are showing aggressive behaviours (including quarrels, fights, displays)
Locomotion	The conspecifics are moving on the ground or in the trees

A naïve observer coded 20% of the complete dataset for reliability by measuring the behaviours independently from the contexts. The two coders reliably classified the same contexts within 3 seconds (i.e., margin of error), Cohen's Kappa ($\kappa = 0.79$). The two coders agreed on the type and number of behaviours, intraclass correlation coefficient (mean ICC(3,1) = 0.59 , SD = 0.24, $p<0.05$).

Statistical analysis

To establish the personality of the chimpanzees of this study, we first examined the contextual and temporal consistency of the behaviours of the focal subject measured in the different contexts and over the two time periods. Then, all behaviours that showed

consistency were subjected to a principal component analysis which allowed us to determine the personality traits that compose the chimpanzees' personality.

Consistency

We used the Generalizability Theory (GT) approach to assess whether the scores of the behaviours generalize across contexts and time (for more information on GT, see Chapter 2; Brennan, 2011; Shavelson & Webb, 1991; Shavelson, Webb, & Rowley, 1989). Although this theory is not frequently used in animal personality studies, it previously demonstrated efficiency and robustness to assess personality in a small sample size of stumptail macaque and zebra finch (Figueredo, Cox, & Rhine, 1995), but also welfare as well as well-being ratings in chimpanzees (Robinson et al., 2017). GT allows to identify multiple sources of systematic and unsystematic errors and estimate the variance components associated with each source of variation.

We used a two-facet crossed design: subject by context by time, where subject represents the object of measurement, and context (two or four levels), as well as time (two levels), represent the facets of generalization. Additionally, we used a one-facet crossed design – subject by time – when a behaviour was observed only in one context across two time periods. For our study, we conducted a Generalizability-study (G-study) on the different facets and their interactions; a coefficient of reliability was then obtained. Two G-studies were carried out considering either the three predominant contexts (feeding and resting) or all four contexts (feeding, affiliative, resting, and solitude). Instead of having a strict cut-off to determine whether the coefficient was acceptable, we treated the coefficient per behaviour as a continuous variable to make use of most of the behaviours we measured in this study. Therefore, the behaviours showed either low, average or high

consistency across context and time. We used a median split to determine which behaviours to retain for further analyses.

Determining the personality structure

Based on the G-study results when considering either two or four contexts, all behaviours that were displayed in more than one context and obtained a coefficient above the median ($Median_{Two\ contexts} = 0.42$; $M_{Four\ contexts} = 0.46$) were averaged over time and across contexts. They were then subjected to a principal component analysis (PCA). To determine the number of components to extract, we used eigenvalue > 0.70 and the scree-plot. By using 0.70 as a threshold for the eigenvalue, we were able to extract four components that explained at least 80% of the variance. Additionally, we used an orthogonal rotation (Varimax). The behaviour loadings exceeding |0.5| were considered as salient (Budaev, 2010). In order to determine which type of rotation was the most appropriate, the analysis was repeated with an oblique rotation (direct Oblimin) which allowed the components to correlate. Whether we considered two (feeding and resting) or four contexts (feeding, grooming-play, resting, solitude), the correlations between the components were relatively low (two-context range: -0.28 – 0.21; four-context range: -0.34 – 0.08). Additionally, they provided identical solutions regarding the behaviour loadings. Therefore, we retained and interpreted the Varimax-rotated components. Subsequently, all behaviours that loaded onto the same component (i.e., personality trait) were summed to create a single personality score for each subject.

Sex and age effects on personality score

We used nonparametric tests to assess whether there was an effect of sex or age on the personality score as these two factors were previously reported to have an impact on personality in chimpanzees (King et al., 2008; Koski, 2011b; Massen et al., 2013). The level of significance was set at 0.05 and the tests were two-tailed.

All analyses were computed using SPSS Statistics 25 (IBM, Chicago, IL). GT analyses were run using the syntax developed by Mushquash and O'Connor (2006). In the syntax, all negative variances were set to zero by default. However, to obtain a more accurate idea of the reliability of the dataset, we ran additional GT analyses where all negative variances were kept in the results (see Table B4 and B5).

3.3 Results

On average, the subjects spent different amounts of time in each of the contexts for each time period (see Table B.2). Aggression represented the lowest duration across all individuals in both 2013 and 2017 (0.21% and 0.51% of the total duration of the dataset, respectively) whereas resting represented the highest in 2013 (36.28% of the total duration of the dataset) and feeding represented the highest in 2017. Across all contexts, the subjects were observed, on average, 42.83 ± 6.20 (mean ± SEM) minutes in 2013 and 97.47 ± 0.91 minutes in 2017. Over the two time periods, the subjects were observed 58.05 ± 2.52 minutes in feeding, 14.63 ± 1.94 minutes in affiliative, 38.94 ± 4.19 minutes in resting, and 19.37 ± 2.99 minutes in solitude contexts.

Additionally, the expression of the behaviours displayed within the four main contexts (feeding, affiliative, resting, solitude) over the two time periods differed across

the individuals (see Table B.3). Figure 3.1 depicts the behavioural expression across subjects within the four contexts per time period.

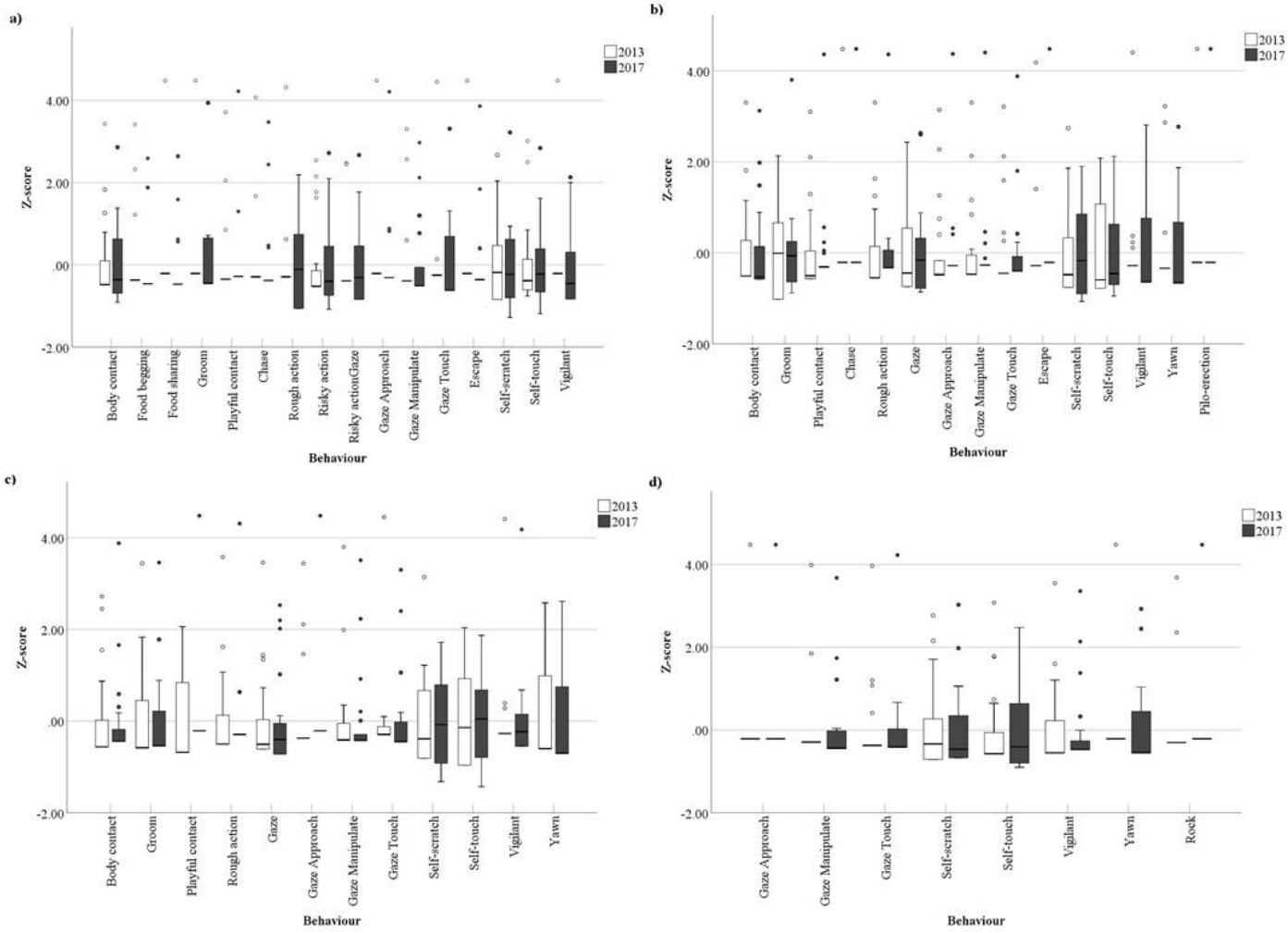


Figure 3.1. Visual representation of the behavioural expression across subjects per time period during a) Feeding, b) Affiliative, c) Resting, and d) Solitude contexts. Data are z-scored within each time period (2013 and 2017). The thick horizontal lines indicate medians; the vertical length of the boxes corresponds to interquartile range; the thin short horizontal lines indicate the minimum and maximum values.

Assessing temporal and contextual consistency

G study results using feeding and resting contexts

We decided to conduct our main analysis using the two main contexts as they together accounted for 69% of the total duration of the dataset (see Table B.2).

Only 11 behaviours were considered for the GT analyses, as the other 11 behaviours were only displayed in one of the two contexts or none of them; among them, one behaviour showed high temporal consistency (Food begging: G = 0.70; see Table 3.6). G-study variance components for each behaviour are presented in Table 3.3 and included 22 subjects. The variance components attributable to subjects were 0% for five of the 11 behaviours (Body contact, Playful contact, Gaze, Gaze Approach, Self-touch) suggesting a lack of inter-individual differences in these specific behaviours. The variance components attributable to subjects for the other six behaviours ranged from 15.20% (Groom) to 34.60% (Gaze Manipulate) suggesting systematic inter-individual differences. Among these six behaviours, the variance attributable to the main effect for measurement contexts, time and the interaction contexts by time accounted for less than 1% of the variance across the behaviours (except for Rough and Self-scratch; see Table 3.3) suggesting that the score for each trait within each subject was consistent across contexts and the two time periods when averaged over subjects and time or over subjects and contexts. The subject by context and subject by time interactions accounted for 0% of the variance for all six behaviours and four behaviours, respectively, suggesting that there was no variability across the individuals in their pattern of contextual and temporal consistency. However, there was some variability between the individuals in their pattern of temporal consistency for Gaze Manipulate (32.20%) and Gaze Touch (4.90%). Finally, the three-way subject by

context by time interaction (residual variance) accounted for a large percentage of variance; this variance reflects different behavioural score per subject across time and contexts as well as other sources of error not systematically included into the G study.

Table 3.3. Estimated variance components for the G-study subject (s) by context (c) by time (t) design and subject by time design per behaviour when considering feeding and resting contexts.

Behaviour	Body contact		Groom		Playful contact		Rough action		Gaze		Gaze Approach		Gaze Manipulate		Gaze Touch		Self-scratch		Self-touch		Vigilant	
Effect	Var	%	Var	%	Var	%	Var	%	Var	%	Var	%	Var	%	Var	%	Var	%	Var	%	Var	%
Subject	0.00	0.00	0.24	15.20	0.00	0.00	0.51	27.10	0.00	0.00	0.00	0.00	0.35	34.60	0.30	29.50	0.67	29.40	0.00	0.00	0.35	26.50
Context	0.01	0.50	0.02	1.00	0.00	0.00	0.02	0.90	0.00	0.00	0.00	0.00	0.00	0.10	0.00	0.00	0.03	1.30	0.00	0.20	0.01	0.40
Time	0.00	0.00	0.01	0.60	0.00	0.00	0.02	1.20	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	1.10	0.00	0.00	0.01	0.60
Subject*Context	0.00	0.00	0.00	0.00	0.01	1.00	0.00	0.00	0.17	15.20	0.01	0.70	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Subject*Time	0.43	36.90	0.00	0.00	0.17	14.90	0.00	0.00	0.27	24.10	0.01	0.70	0.33	32.20	0.05	4.90	0.00	0.00	0.37	32.80	0.00	0.00
Context*Time	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Subject*Context*Time	0.72	62.50	1.30	83.30	0.96	84.10	1.34	70.90	0.69	60.70	1.07	98.50	0.34	33.10	0.66	65.50	1.56	68.30	0.75	67.00	0.96	72.40

Var: Variance estimate was calculated using the mean squares; High values suggest that there is a large effect of the variable (see section 2.2.3, Table 2.1). %: Percentage of the total variance for each variance estimate was calculated by dividing each variance estimate by the total variance.

As the variance from the subject was 0 for five of the 11 behaviours, their G-coefficients were equal to 0 (see Table 3.4), suggesting that these behaviours were not consistent across contexts and time. Regarding the other six behaviours, the G-coefficients varied from 0.42 to 0.63 suggesting consistency across both context and time.

Following Figueredo and colleagues (1995), we constructed the G-coefficients for context and time separately for each of the six behaviours (see Table 3.4). For context, the G-coefficients were equal to 1, indicating high reliability across feeding and resting contexts when averaged over subjects and time. Concerning time, the G-coefficients varied from 0.52 to 1, indicating relatively high reliability overall over time when averaged over subjects and contexts.

Table 3.4. Overview of the G-coefficients for subject by context by time design as well as for context and time separately for each behaviour when considering the feeding and resting contexts.

Behaviour	G-coefficient	G-coefficient for context	G-coefficient for time
Body contact	0.00	-	-
Groom	0.42	1.00	1.00
Playful contact	0.00	-	-
Rough action	0.60	1.00	1.00
Gaze	0.00	-	-
Gaze Approach	0.00	-	-
Gaze Manipulate	0.59	1.00	0.52
Gaze Touch	0.61	1.00	0.86
Self-scratch	0.63	1.00	1.00
Self-touch	0.00	-	-
Vigilant	0.59	1.00	1.00

The additional GT analysis which replaced the variances of 0 with the original negative values revealed a similar pattern for the different behaviours, although the values were larger overall (see Table B.4). The variance components attributable to

subjects that were 0% for Body Contact, Playful contact, Groom, Gaze Approach, Self-touch and Self-scratch had negative values which could suggest a problem in the dataset; i.e., low amount of data point per subject which may lead to very little variation between individuals. The variance attributable to the main effect for measurement contexts were all above 0 for the other behaviours, suggesting that they were consistent across contexts when averaged over subjects and time. The variance attributable to the main effect for time had a small negative value (less than - 0.10) for some of the behaviours (Gaze Manipulate and Gaze Touch); as recommended in the literature (see Chapter 2 for details), these values can be set to zero, suggesting consistency over time when averaged over subjects and contexts. With regard to the interactions between the different facets (context and time), for some of the behaviours, the negative values had a bigger magnitude than for the main effect measurements, ranging from - 0.478 to - 0. 123; this suggests that we need to consider more carefully the validity of the interactions and their meaning in regards to the model due to a low amount of data per individual. Finally, the variance for the three-way interactions is relatively high for some of the behaviours as a result of the presence of negative values for the main effect and interaction measurements. All behaviours that were originally selected for the PCA would also be selected if using this second set of GT analysis.

Although it is important to be cautious when considering the consistency aspect of these behaviours for the main effect (time) and the interactions, it suggests a strong tendency of consistency, and it is likely that the pattern of consistency would be more substantial if there were more data per individuals.

G study results using feeding, affiliative, resting and solitude contexts

We decided to re-run the analysis using all contexts that had a total duration across the two time periods representing at least 10% of the overall total duration of the dataset (see Table B.2).

Three behaviours (Risky approach, Display, Throw) were discarded from the analyses as they were only expressed either in one time period or one context. G-study variance components for each behaviour are presented in Table 3.5 and included 22 chimpanzees. Five behaviours were displayed only in one context across the two periods (Food begging, Food sharing, Risky action, Pilo-erection, Rock); none of them showed temporal consistency (G = 0; see Table 3.6), except for Food begging as mentioned above. Among the 14 remaining behaviours, six of them (Escape, Playful contact, Chase, Gaze Approach, Gaze, Self-touch) had a variance of 0% suggesting a lack of inter-individual differences. The variance components attributable to subjects for the other eight behaviours ranged from 2.50% (Body contact) to 36.30% (Gaze Touch) suggesting inter-individual differences in their scores. The variance attributable to the main effect for measurement contexts, time and the interaction contexts by time accounted for less than 1% of the variance across the eight behaviours suggesting that the score for each behaviour within each subject was consistent across contexts and time periods when averaged over subjects and time or over subjects and contexts. The subject by context interaction accounted for 0% of variance for all behaviours but Self-scratch (4.30%), and the subject by time interaction accounted for 0% for all eight behaviours, indicating very little, if not, variability across the individuals in their pattern of contextual and temporal consistency for the eight behaviours. As in the previous G-study analysis, the three-way subject by context by time interaction (residual variance) accounted for a large percentage of variance.

Table 3.5. Estimated variance components for the G-study subject (s) by context (c) by time (t) design and subject by time design per behaviour when considering the feeding (F), affiliative (GP), resting (R) and solitude (S) contexts.

Context	F-GP-R		F-GP-R		F-GP-R		F-GP		F-GP-R		F-GP-R		F-GP-R-S	
Behaviour	Body contact		Groom		Playful contact		Chase		Rough action		Gaze		Gaze Approach	
Effect	Var	%	Var	%	Var	%	Var	%	Var	%	Var	%	Var	%
Subject	0.03	2.50	0.25	18.00	0.00	0.00	0.00	0.00	0.29	21.70	0.00	0.00	0.00	0.00
Context	0.01	0.60	0.01	1.00	0.00	0.00	0.00	0.00	0.01	0.80	0.00	0.00	0.00	0.00
Time	0.00	0.00	0.00	0.10	0.00	0.00	0.00	0.00	0.00	0.20	0.00	0.00	0.00	0.00
Subject*Context	0.00	0.00	0.00	0.00	0.18	16.10	0.81	75.40	0.00	0.00	0.23	19.90	0.02	1.90
Subject*Time	0.11	9.30	0.00	0.00	0.05	4.10	0.00	0.00	0.00	0.00	0.19	16.70	0.16	14.90
Context*Time	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Subject*Context*Time	1.01	87.70	1.10	80.80	0.88	79.70	0.26	24.60	1.03	77.20	0.72	63.40	0.89	83.20
Context	F-GP-R-S		F-GP-R-S		F-GP		F-GP-R-S		F-GP-R-S		F-GP-R-S		GP-R-S	
Behaviour	Gaze Manipulate		Gaze Touch		Escape		Self-scratch		Self-touch		Vigilant		Yawn	
Effect	Var	%	Var	%	Var	%	Var	%	Var	%	Var	%	Var	%
Subject	0.28	28.00	0.40	36.30	0.00	0.00	0.16	15.30	0.00	0.00	0.30	27.80	0.34	28.70
Context	0.00	0.00	0.00	0.30	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.40
Time	0.00	0.00	0.00	0.10	0.00	0.00	0.00	0.20	0.00	0.00	0.00	0.40	0.00	0.20
Subject*Context	0.03	3.10	0.00	0.00	0.01	0.50	0.05	4.30	0.10	10.20	0.00	0.00	0.00	0.00
Subject*Time	0.10	9.60	0.00	0.00	0.19	17.90	0.00	0.00	0.21	21.00	0.00	0.00	0.00	0.00
Context*Time	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Subject*Context*Time	0.59	59.30	0.70	63.20	0.88	81.60	0.83	80.20	0.69	68.90	0.79	71.80	0.84	70.60

Context	F		F		F		GP		S	
Behaviour	Food begging		Food sharing		Risky action		Pilo-erection		Rock	
Effect	Var	%	Var	%	Var	%	Var	%	Var	%
Subject	0.55	54.20	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Time	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Subject*Time	0.46	45.80	1.10	100.00	1.05	100.00	1.05	100.00	1.07	100.00

Among the eight behaviours that showed inter-individual variance across both contexts and time, the G-coefficients varied from 0.34 to 0.82 indicating low to high-reliability scores overall (see Table 3.6).

As previously done, we constructed the G-coefficients for context and time separately for each of the eight behaviours (see Table 3.6). For context, the G-coefficients ranged from 0.78 to 1, indicating high reliability across feeding, affiliative, resting and solitude contexts when averaged over subjects and time. Concerning time, the G-coefficients varied from 0.21 to 1, indicating low to high reliability over time when averaged over subjects and contexts.

Table 3.6. Overview of the G-coefficients per behaviour for subject by context by time design and for subject by time design when considering the feeding (F), affiliative (GP), resting (R) and solitude (S) contexts. The table depicts the G-coefficients for both context and time separately.

Context	Behaviour	G-coefficient	G-coefficient for context	G-coefficient for time
F-GP-R	Body contact	0.34	1.00	0.21
F-GP-R	Groom	0.57	1.00	1.00
F-GP-R	Playful contact	0.00	-	-
F-GP	Chase	0.00	-	-
F-GP-R	Rough action	0.63	1.00	1.00
F-GP-R	Gaze	0.00	-	-
F-GP-R-S	Gaze Approach	0.00	-	-
F-GP-R-S	Gaze Manipulate	0.68	0.90	0.74
F-GP-R-S	Gaze Touch	0.82	1.00	1.00
F-GP	Escape	0.00	-	-
F-GP-R-S	Self-scratch	0.58	0.78	1.00
F-GP-R-S	Self-touch	0.00	-	-
F-GP-R-S	Vigilant	0.76	1.00	1.00
GP-R-S	Yawn	0.71	1.00	1.00
F	Food begging	0.70	-	-
F	Food sharing	0.00	-	-
F	Risky action	0.00	-	-
GP	Pilo-erection	0.00	-	-
S	Rock	0.00	-	-

As previously found for the analysis involving only two contexts, the additional GT analysis revealed a similar pattern for the different behaviours, although the values were larger overall (see Table B.5). The variance components attributable to subjects that were 0% for Food sharing, Risky action, Pilo-erection, Rock, Escape, Playful contact, Chase, Gaze, Gaze Approach and Self-touch had negative values which could suggest a problem in the dataset; i.e., low amount of data point per subject which may lead to very little variation between individuals. The variance attributable to the main effect for measurement contexts and time had a small negative value (less than - 0.10) for some of the behaviours (Body contact, Gaze Manipulate, and Self-scratch). With regard to the interactions between the different facets (context and time), for some of the behaviours (Body contact, Groom, Rough action, and Yawning), the negative values had a bigger magnitude than for the main effect measurements, ranging from - 0.311 to - 0.114. Finally, the variance for the three-way interactions is relatively high for some of the behavioural actions as a result of the presence of negative values for the main effect and interaction measurements. All behaviours that were originally selected for the PCA would also be selected if using this second set of GT analysis.

As mentioned in the previous analysis, the interpretation of the consistency for some of the behaviours has to be considered carefully here; however, a bigger dataset would likely provide stronger patterns of consistency for most of the behaviours.

Personality structure

Based on GT including feeding and resting

We decided to keep all behaviours that had a G-coefficient above the median (M = 0.42); hence, 6 behaviours were retained for the PCA. The Kaiser–Meyer–Olkin (KMO = 0.568) and Bartlett's Test of Sphericity ($p < 0.01$) indicated the suitability of the data. We extracted four components which explained 89.8% of the variance (see Table 3.7). The first component explained 33.5% of the variance and had positive loadings of two behaviours that related to explorative behaviours; this component was labelled "exploratiVeness". The second component explained 21.9% of the variance and had positive loadings of two behaviours related to anxious behaviours; this component was therefore labelled "anxiety". The last two components accounted for 17.6% and 16.8% of the variance, respectively. Only one behaviour positively loaded on each of the two components; one of the two behaviours related to bold actions whereas the second one related to social actions. We labelled them "boldness" and "sociability", respectively.

Table 3.7. Behaviour loadings after Varimax rotation when considering the feeding and resting contexts.

Behaviour	ExploratiVeness	Anxiety	Boldness	Sociability
Groom				0.978
Rough action			0.983	
Gaze Manipulate	0.943			
Gaze Touch	0.951			
Self-scratch		0.816		
Vigilant		0.790		

Only loadings above |0.5| are reported.

Based on GT including feeding, affiliative, resting and solitude

We decided to keep all behaviours that had a G-coefficient above the median (M = 0.46); hence, seven behaviours were retained for the PCA. The Kaiser–Meyer–Olkin (KMO = 0.545) and Bartlett's Test of Sphericity ($p < 0.001$) indicated the suitability of the data. Overall, when considering the four contexts, we obtained similar loadings compared to the PCA that included feeding and resting. We extracted four components which explained 86.7% of the variance (see Table 3.8). The first component explained 30.3% of the variance and had positive loadings of two behaviours that related to explorative actions; this component was labelled "exploratiVeness". The second component explained 21.9% of the variance and had positive loadings of two behaviours; one behaviour was related to social actions and the other one was associated with anxiety-related actions. This component was therefore labelled "sociability". The last two components accounted for 19.9% and 14.6% of the variance, respectively. The first one had positive loadings of two behaviours that were related to anxious actions; we labelled this component "anxiety".

Only one behaviour positively loaded onto the last component and was associated with bold actions; we labelled the component "boldness".

Table 3.8. Behaviour loadings after Varimax rotation when considering the feeding, affiliative, resting and solitude contexts.

Behaviour	Exploratiyeness	Sociability	Anxiety	Boldness
Groom		0.768		
Rough action				0.989
Gaze Manipulate	0.957			
Gaze Touch	0.968			
Self-scratch			0.908	
Vigilant			0.679	
Yawn		0.906		

Only loadings above |0.5| are reported.

Sex and age effects on personality

Based on the PCA solution that considered feeding and resting contexts, we summed the behaviour that loaded onto the same component, creating, therefore, four personality traits. There was no effect of sex on any of the personality trait scores (see Table B.6). However, there was a negative significant correlation between the age of the subject and their boldness score (Spearman's correlation coefficient: $r_s = -0.43$, $p < 0.05$; see Table B.7).

3.4 Discussion

This study examined 22 behaviours expressed by semi-wild chimpanzees during naturally occurring contexts in the years 2013 and 2017. These behaviours were primarily expressed in two naturally occurring contexts, that is, feeding and resting, but were also displayed in affiliative and solitude contexts. Among these 22 behaviours, six of them were considered

reliable across feeding and resting over time. These six behaviours reflected four different personality traits: exploratiVeness, anxiety, boldness, and sociability. A similar personality structure (i.e., the four traits were constituted of similar behaviours) was observed when considering all four contexts (feeding, affiliative, resting and solitude), implying that the personality profile of these chimpanzees is relatively stable. Considering such a range of contexts helped contribute to a better description of each underlying personality trait. The behaviours measured were expressed differently across contexts, implying that some contexts play a more important role than others in terms of consistency depending on the behaviour examined. Previous research that reported contextual consistency in nonhuman primates was only measured across 2-6 experimental situations (e.g., (Šlipogor et al., 2016; Uher et al., 2008), that were arguably similar to the extent that they could be categorized as belonging to the same singular context. Here, multiple distinct contexts were used to cover various behaviours which reflected four personality traits.

In this study, contextual consistency was even found between contexts that differ on the functional, affective, arousal as well as social level. For instance, exploratiVeness-related behaviours were consistent across both solitary and social contexts, and sociability-related behaviours were consistent across feeding, affiliative, and resting contexts. This consistency pattern was even stronger when considering only the two main contexts (feeding and resting), namely the ones that represented over 25% of the overall dataset. Thus, using such an approach that measures personality in a natural setting allows us to measure and find a range of distinctive contexts. This way, we can capture more the complexity of a personality trait in the daily life of a species.

Past research that examined contextual and temporal consistency revealed consistency for nonhuman primate personality and was observed for a range of personality traits, including boldness, explorativeness, anxiety (Kutsukake et al., 2012; Massen et al., 2013; Uher et al., 2008) as well as sociability (Koski, 2011b). In our study, each measured behaviour that reflected the different personality traits showed consistency in some contexts and not in others, showed no consistency over time or low to high level of consistency. These different patterns among the behaviours may be explained in that the contexts and/or time periods affect to some extent the expression of the personality-linked behaviours. Such effects may result in higher level of consistency between some contexts over some periods of time as previously observed in an experimental study in common marmosets (Koski & Burkart, 2015). It is possible that some behaviours may not show contextual and/or temporal consistency – key criteria that are often associated with personality assessment (Gosling, 2008). Yet, this lack of consistency does not necessarily mean that these behaviours do not play a significant role in the personality trait of an individual. Capturing the wide range of distinctive contexts, by including the ones that elicit different occurrences of personality-based behaviours, is highly important for understanding the personality trait of interest. This way would be a valuable step to take in future studies on nonhuman primate personality. In addition, as done in human personality research (Funder & Colvin, 1991; Wu & Clark, 2003) and a few nonhuman primate studies (Tomassetti et al., 2019; Uher, Addessi, et al., 2013; Uher & Asendorpf, 2008; Uher et al., 2008; Uher & Visalberghi, 2016), future personality research on nonhuman primates may perhaps benefit from assessing consistency on the trait level rather than the behaviour level. By aggregating all behaviours that are likely to reflect conceptually the same trait may capture more the variance of expression of a personality trait.

While animal personality research has been often examined and discussed as being consistent (Gosling, 2001, 2008), we found a notable number of exceptions in the data. The GT analysis allowed us to explore how much flexibility there was in terms of both contextual and temporal consistency in chimpanzees living in natural settings. The results revealed that there was some variability in terms of both contextual and temporal consistency of the different behaviours that reflected the four personality traits across the subjects. Individual primates may adjust to specific situations they could find themselves in (Dingemanse & Wolf, 2013), which they benefit from fitness-wise (Wolf & Weissing, 2012). Different rearing environments (Bard & Gardner, 1996; Martin, 2005) or different roles played by members within a social group (Foerster et al., 2016), for instance, may shape individuals' behaviours throughout their lives. These factors could possibly result in variations in the expression of their personality traits over time and across contexts, and in inter-individual differences, contributing therefore to the expression of different profiles of personality traits (Furr & Funder, 2004; Uher et al., 2008). Future research would benefit from focusing on both the behaviour/personality trait and the personality profile of an individual. Such an approach would provide more details in terms of inter-individual variability in personality consistency.

Different factors could influence the personality consistency, such as the audience surrounding the focal subject. Chimpanzees are known to be able to adjust their behaviours or decisions according to their social surrounding when it comes to social play (Flack et al., 2004) or grooming (Mielke et al., 2018). It is possible that similar patterns are observed with behaviours that reflect an individual's personality, as the subjects may be keen on adjusting their behaviours for their own benefits. However, it is unlikely to have had an impact on our data as we found consistency across both social and solitary contexts.

Regarding sex and age effects, previous research reported conflicting outcomes. Some studies showed that males were more anxious, active or dominant than females (Dutton, 2008; Koski, 2011b; Pederson et al., 2005; Weiss & King, 2015) or found that younger chimpanzees were bolder than older individuals (Massen et al., 2013), whereas other studies did not find any effect of sex (Massen et al., 2013) or age (Herrmann, Hare, Cissewski, & Tomasello, 2011). Our study revealed no effect of the sex of the subjects on their personality scores, but younger chimpanzees were significantly bolder than their counterparts, which supports previous findings (Massen et al., 2013).

It is, however, important to remain cautious in the interpretation of the results, as our study also presents some limitations. Using behavioural observations in natural setting has great advantages in terms of ecological validity (Freeman et al., 2011) but also has some drawback in terms of the amount of data obtained per subject. The subjects of our study had fewer behavioural observations (i.e., about 53h of recordings). The subjects of our study had fewer behavioural observations (i.e., about 53h of recordings) in comparison to other personality primate studies (e.g., Eckardt et al., 2015; Konečná et al., 2008; Koski, 2011b; Neumann et al., 2013). However, it is not uncommon to see behavioural studies with fewer hours of recordings per subject (Baker et al., 2015; Maestripieri, 2000; Uher et al., 2008), especially experimental studies, in which the subjects are observed mostly twice in each experimental situations to assess time consistency (Koski & Burkart, 2015; Massen et al., 2013; Šlipogor et al., 2016; Tkaczynski et al., 2018). It is possible that the amount of data per individual was not sufficient to provide a clear picture of their personality profile. In this study, we used a very detailed coding approach allowing us to code independently multiple distinct contexts and various behaviours in order to determine inter-individual differences. In the analyses, we tried to adopt a cautious approach by focusing, first, on the

two predominant contexts where most of the data were available, and we, then, run a second set of analysis by considering all relevant contexts. Furthermore, we used a robust statistical framework – Generalizability Theory – which has previously demonstrated its efficiency and validity to assess reliably personality with small sample sizes (Figueredo et al., 1995), avoid biased estimation in sampling and scoring decisions (Hernández-Lloreda & Colmenares, 2006) and assess the reliability of welfare and wellbeing ratings (Robinson et al., 2017). This theory allowed us to separate the different sources of errors coming potentially from context, time and their interactions. However, by looking at the interactions, we, therefore, reduced even more our dataset available which resulted occasionally in negative variances for some of the behaviours. Such negative values have been reported to indicate a sampling error (Shavelson & Webb, 1991). Yet, in case of small negative variances (very close to 0), it is widely recommended to set the variances to 0 (Brennan, 2001), which, to some extent, may, then, suggest a lack of variation across context and/or time (i.e., consistency). Although the findings presented in this study must be carefully considered, it is likely that if the individuals had more data points per context and per time period, we would obtain stronger patterns of behavioural consistency. To our knowledge, the current study provides first evidence for contextual consistency of personality traits across multiple distinctive contexts in a nonhuman primate species when observed outside a captive, zoo-like setting.

To conclude, our findings showed that the consistency of sociability, boldness, explorativeness and anxiety can be investigated across multiple distinct naturally occurring contexts. The naturalistic approach of this study allowed us to cover a wide range of contexts, revealing that some have a greater impact on the individual's personality consistency (e.g., feeding, resting) than others (e.g., solitude, affiliative), and consequently

contributed to a better understanding of the complexity of a personality trait. Combining such an approach with ratings and/or experimental observations could provide a detailed description of the personality structure of nonhuman primate species. Our study provides new insight into animal personality research, as depending on the personality trait examined, consistency may occur only across some contexts and over time. This work suggests that personality consistency is more malleable than initially assumed, once focusing on the significance of the contexts. It is important to consider such a range of contexts and behaviours that are frequently expressed across nonhuman primate species. In doing so, we can highlight variations, and thus, try to identify what selective pressures may have played a role in the emergence of some personality traits in humans (Buss, 1988; Nettle, 2006).

4. Assessing personality in sanctuary chimpanzees (Pan troglodytes): Comparing naturalistic, rating and experimental approaches

Abstract

Three main methods have been developed to measure animal personality: ratings, experimental assays and naturalistic observations. Very few studies used a combination of methods to measure personality in primates, and none of them systematically measured the same trait. A lack of positive correlation between methods has often been associated with methodological discrepancies. The aim of this study was to systematically compare the three methods, with each of them measuring the same trait (i.e., sociability, boldness, exploratiVeness, anxiety) in 24 sanctuary chimpanzees. Naturalistic observations consisted of ecologically-relevant contexts and a wide range of behaviours reflecting the traits of interest. A questionnaire that combined ratings of adjective-behaviour items to measure the four traits was developed, and two experiments to measure two of these traits, boldness and exploratiVeness, were conducted. Our study revealed mix findings in terms of coherence between methods. Ratings positively correlated with experimental observations for exploratiVeness and naturalistic observations correlated with rating methods for sociability, boldness and anxiety. Negative trends were found for anxiety when comparing ratings with naturalistic observations, and naturalistic observations with the experimental observations. Overall, high coherence was shown when considering the context in which the personality trait was measured. To capture more diversity of expression of the different personality traits, it is important for future research to consider more the context.

4.1 Introduction

Personality is characterised by multiple traits, dynamically organised (Allport, 1961). It can be a complex concept to fully understand as individuals can express their different personality traits using diverse behaviours in distinct daily contexts or situations throughout life (Eaton & Funder, 2001; Mehl et al., 2006). For instance, an individual's personality can be measured in daily interactions (e.g., dyadic, group, phone or self-talk conversations), various activities (e.g., watching TV, attending class) or locations (e.g., home, restaurants) (Mehl et al., 2006). This diversity of personality expression suggests perhaps that the same personality trait may have multiple facets of expression depending on the situation experienced. For instance, an individual's aggressiveness is likely to be expressed differently when watching a national game on the TV at home and when being involved in a fight on a Saturday night. Personality affects individuals' everyday life, decision making, and their relationships with others (Furr, 2009). These inter-individual and intra-individual differences are likely to result in differences in the expression of their behaviour-based traits (e.g., being bold or shy), making individuals unique (Allport, 1937). However, somewhat paradoxically, personality traits are reported to be stable over time and/or across situations (Allport & Odbert, 1936; Furr & Funder, 2004; Weisbuch et al., 2010).

In nonhuman primates, both temporal and contextual consistency has been shown in various primate species ranging from the species most distantly related to humans – for instance, the grey mouse lemur (*Microcebus murinus*) (Dammhahn, 2012; Dammhahn & Almeling, 2012) – to those most closely related to humans – chimpanzees (Koski, 2011b; Massen et al., 2013). Animal personality has been often perceived as rigid in terms of its

definition (Gosling, 2001, 2008), where consistency is a key criterion. However, it is important to note that although some behaviours may not show consistency over time or across contexts (Chapter 3), they may still play an important role in the personality structure of an individual. Consequently, individuals may show different patterns of consistency in different traits across different contexts and over time (Furr & Funder, 2004; Uher et al., 2008).

Three approaches have been developed to assess nonhuman primates' personality: ratings, naturalistic and experimental observations (King & Figueredo, 1997; Koski, 2011b; Kutsukake et al., 2012; Weiss et al., 2007). The rating approach requires familiar observers or carers to rate the individuals' personality using a series of adjective or behaviour-descriptor verb items (see section 2.1.1; Bard & Gardner, 1996; Clay et al., 2015; King & Landau, 2003; Stevenson-Hinde & Zunz, 1978; Uher & Asendorpf, 2008; Uher & Visalberghi, 2016; Weiss & King, 2015). It can quickly obtain personality data in a large sample in both captive and wild settings (King et al., 2005; Tkaczynski et al., 2018; Weiss et al., 2007, 2017) and can predict behaviours when observed in a natural setting (Capitanio, 1999; Freeman et al., 2013; Pederson et al., 2005). For instance, dominance was positively correlated with agonistic behaviours, such as aggressive display, in captive chimpanzees (Pederson et al., 2005). The adjective items used in questionnaires represent often a clear reflection of the targeted personality traits (e.g., bold or cheeky to target boldness), and are, therefore, assumed to reflect the general representation of a personality trait.

However, the rating approach garnered criticisms in term of its subjectivity and its anthropomorphic aspect (Uher, 2008a). On one hand, the raters' judgements can be influenced by major events (e.g., big fight) occurring in the life of the focal subject or by

another rater who witnessed an unusual event a few weeks before the rating took place, for instance (Freeman et al., 2011; Uher, 2008a). On the other hand, the adjectives selected to measure animal personality may not be always fully adapted to the studied species (e.g., Adams et al., 2015; Eckardt et al., 2015; King & Figueredo, 1997). For instance, adjectives such as "jealousy" or "cool" are probably more relevant to describe human personality (Goldberg, 1990) rather than animals (Eckardt et al., 2015). Uher and colleagues (2008; 2016; 2013) developed a different rating approach, which combines items with behaviour-descriptive verbs and trait-adjectives, to measure various personality traits in captive primate species. Unlike the adjective items, these behavioural descriptors provide a more descriptive assessment of the personality trait targeted as the personality trait is associated with a specific situation (Uher, 2008a, 2008b, 2011). This narrower perspective is, however, likely to target only a facet of the measured personality trait compared to the adjective items.

Unlike the rating approach, the behavioural observational approach is objective and relies only on the coding of behaviours expressed by the subjects (Freeman & Gosling, 2010; Koski, 2011a). Experimental assays were designed to measure specific personality traits in controlled environments and elicit specific behavioural reactions towards a stimulus across subjects (see section 2.1.2; Freeman & Gosling, 2010; Freeman et al., 2011). For instance, in a novel object or food test, the subjects are expected to display various degrees of exploration or avoidance such as gazing, touching, handling or moving away from the stimulus (Gosling, 2001; Massen et al., 2013; Réale et al., 2007). Yet, this experimental approach may lack some ecological validity, as the subjects are tested in an environment that rarely reflects natural events. In addition, this approach focuses only on a limited number of behaviours displayed in a single everyday-like situation (e.g., snake test

to measure boldness in a predator-like situation: Koski & Burkart, 2015; Massen et al., 2013). Spontaneous behaviours that could possibly emphasise subtle inter-individual variations are thus ignored. Such a method is likely to target only a specific facet of a personality trait, as the experimental observations are closely associated with a specific situation.

In contrast with the other two methods, the naturalistic observational approach is an objective method of high ecological validity (see section 2.1.3). It can be applied in multiple distinct contexts (e.g., play, feeding, resting) providing a more detailed description of individuals' personality (Gosling, 2008; Mehta & Gosling, 2008; Uher, 2008a) as opposed to more specific experimental situations. Observing the behaviour in the natural course of the individuals' lives in a natural setting is likely to cover a wide range of relevant contexts, revealing possibly the impact of some contexts (e.g., affiliative) over others (e.g., resting, solitary) (see Chapter 3), and can be applied without manipulation of the subjects. Furthermore, it is likely to contribute to a better understanding of the complexity of a personality trait as it is likely to cover multiple facets and highlight potential inter-individual differences in a wide range of natural contexts (see Chapter 3).

However, as the other two methods, the naturalistic approach has some drawbacks (see section 2.1.3; Freeman & Gosling, 2010; Freeman et al., 2011). This approach cannot control for fluctuations coming from the environment which may affect the behavioural responses of the subjects. In addition, as the observations occur in a natural setting, the subjects may decide to leave the visible area at any moment during the observation period. Finally, this approach can be time-consuming in terms of behavioural coding as the subjects have to be observed on multiple occasions across different contexts, so the researcher can capture the diversity of expression of the subjects' personality.

Naturalistic observations have been combined with the rating or experimental approach only to validate new scale or to assess convergent and discriminant validity of the personality traits (Eckardt et al., 2015; Pederson et al., 2005; Weiss et al., 2007, 2015). Various findings have been reported in regard to the coherence between behavioural observations and the personality dimensions obtained with the rating approach (Konečná et al., 2008, 2012; Pederson et al., 2005; Schaefer & Steklis, 2014). For instance, Uher and Asendorpf (2008) found higher coherence between experimental behavioural observations and behaviour-descriptive verb items than between experimental behavioural observations and adjective items; however, both formats of items showed coherence in their measurements (Uher & Asendorpf, 2008). An incongruence between methods to measure diverse personality traits has been often described as a result of methodological differences (Tkaczynski et al., 2018) or the effect of the rater's subjective ratings (Uher, 2008a). However, it is possible that lower coherence between methods may reflect more the complexity of examining personality. Instead, these different methods may capture different facets of a personality trait (Carter et al., 2012a). Therefore, including different approaches, each of which measuring the same personality trait, would help us understand better the full expression of a trait, and, therefore, obtain a more accurate picture of individual's personality traits.

To our knowledge, only seven studies systematically used and compared different methods to assess the temporal and/or contextual consistency of different personality traits (Capuchin *sp.*: Uher & Visalberghi, 2016; *Macaca fascicularis*: Uher, Werner, et al., 2013; *Macaca nigra*: Neumann et al., 2013; *M. sylvanus*: Tkaczynski et al., 2018; *Papio ursinus*: Carter et al., 2012a; four great ape species: Uher & Asendorpf, 2008; Chimpanzees: Freeman et al., 2013). They reported positive correlations between the rating and

experimental or naturalistic approaches for various personality traits including sociability, boldness, explorativeness or anxiety (Carter et al., 2012a; Freeman & Gosling, 2010; Tkaczynski et al., 2018; Uher et al., 2008; Uher & Visalberghi, 2016; Uher, Werner, et al., 2013). However, some of these studies also found a lack of coherence between any of the personality traits measured (Neumann et al., 2013; Tkaczynski et al., 2018). Among these seven studies, only one used all three approaches to determine their degree of comparability and their applications in wild settings in wild Barbary macaques (Tkaczynski et al., 2018). Although all personality traits measured were not systematically assessed across all three methods of personality assessment, they found positive correlations between naturalistic observations and ratings for sociability-related traits.

Here, we compared three methods to measure four personality traits, i.e., sociability, boldness, explorativeness and anxiety, in sanctuary chimpanzees. First, we used a naturalistic approach that targeted diverse behaviours that reflected the four traits as well as they occur across multiple distinct contexts (Chapter 3). Additionally, based on Uher and colleagues' work (Uher & Asendorpf, 2008; Uher & Visalberghi, 2016; Uher, Werner, et al., 2013), we developed a short questionnaire that asked keepers to rate the four traits of interest using a combination of adjective and behaviour-descriptive verb items. Finally, based on previous experimental studies in nonhuman primates (Koski & Burkart, 2015; Massen et al., 2013; Seyfarth et al., 2012; Šlipogor et al., 2016), we conducted two experiments – a snake test and a novel object test – to target boldness and explorativeness, respectively. Although we were interested in four specific personality traits, we used a traditional approach in this study (see section 2.3.1), that is, only repeatable behaviours or items were used to determine the structure of each personality trait.

Each of the four personality traits of interest have been previously reported to be measurable using each of the three methods of assessment (see Chapter 3; Anestis, 2005; Freeman et al., 2013; Koski, 2011b; Massen et al., 2013; Uher et al., 2008; Weiss et al., 2007). In this study, we determined whether each method measured the same trait, more specifically the same facet of expression of the personality trait. Based on past research (e.g., Carter et al., 2012a; Tkaczynski et al., 2018; Uher & Asendorpf, 2008; Uher & Visalberghi, 2016), we hypothesise that the three approaches show similarities in their assessment of the personality traits. More specifically, when compared on the trait level, we expect to see high coherence 1) between ratings (adjective item) and naturalistic observations, 2) between ratings (behaviour-descriptive verb item) and experimental observations, and 3) between the two formats of items. However, we do not expect any specific relationship between experimental and naturalistic observations due to their respective particularity (close vs broad perspective; see section 2.1.4). When compared on the behaviour level, we expect to see high coherence between naturalistic observations and ratings (behaviour-descriptive verb item).

4.2 Material and Methods

Subjects and study site

All subjects included in this study lived at Chimfunshi Wildlife Orphanage (Zambia). Depending on the method of assessment, the number of subjects varied across methods (see Table C.1). For the naturalistic approach, 23 subjects (8 females) were included, with an age range of 6 and 36 years old (mean ± SD: 19.30 ± 9.84). For the rating approach, 24 chimpanzees (9 females) were included (age range: 6-36 years; 18.96 ± 9.77). And for the experimental approach, 10 (3 females) and 13 (6 females) chimpanzees were tested in the

snake and novel object tests, respectively; the age ranged from 9 to 35 years old (19.57 ± 9.87). All individuals observed in the experimental setting were also observed in the natural setting, and all individuals observed in the natural setting were rated by the keepers.

The subjects were members of two stable, multi-male multi-female colonies, showing natural fission-fusion dynamics. The larger colony (Colony 1) comprised 24 chimpanzees in a 77-hectare enclosure, and the second colony (Colony 4) included 11 chimpanzees in a 25-hectare enclosure. Each enclosure contained naturally developed fruit groves, grasslands and forests in the miombo woodland, as well as an indoor area (used for midday feeding or medical check-up). Approximately 1.5 kilometres separated the two colonies preventing the subjects to observe chimpanzees from the other colony. The chimpanzees were provisioned with food (e.g., local fruits and vegetables) inside around noon and outside in the afternoon each day, and they could also forage in the forest. Water was always available outside via a water fountain. Further information regarding the formation history of the two colonies as well as their genetic background is provided in Chapter 3.

Data collection and coding

Naturalistic approach

Video-recordings were collected using focal animal sampling (Altmann, 1974); the subjects were followed for 4 min (2013) or 5 min (2017), once or twice a day during the morning (from 7:30 to 11:30 am) and/or afternoon (from 1 to 5:30 pm) sessions. Recording occurred over two field seasons, between June and September 2013 and between May and September 2017. Prior to each recording session, the order of focal animal sampling was randomised to avoid any bias towards the same individual and provide a balance between

time periods and contexts. Thus, all subjects were observed at different time throughout the day and in different contexts (social and non-social). Each conspecific was individually identified by the observer as soon as they were present within 10 meters of the focal subject. A total of 53 hours of recording was collected for this study, with approximately 2.5 hours per individual (Mean ± SD = 2.43 ± 0.50 hours). For further information regarding the representativeness of the amount of collected data per individual per time period across contexts and per individual per context over the two time periods, see Chapter 3.

Behaviours. Twenty-two behaviours were coded based on previous personality and behavioural studies in nonhuman primates (e.g., Baker & Aureli, 1997; Massen et al., 2013; Nishida et al., 1999; Pederson et al., 2005; Uher et al., 2008). Each of these behaviours which reflected four common personality traits: sociability, boldness, exploratiVeness and anxiety (see Table 3.1 for definitions). These behaviours were coded independently from the contexts. The same action was not considered as part of the same series if there was a break of at least 5 seconds. This approach allowed to ensure the independence of the occurrences and avoid a possible inflated estimation of inter-individual differences. For each type of behaviour, we computed frequency per hour of context. Each subject obtained one single score per behaviour for each context, and each score was standardized as a z-score.

Context. The context was coded frame by frame and according to the presence and activity of the conspecifics (not of the subject) present within 10 metres of the focal. To be considered, a context had to be displayed by at least half of the conspecifics surrounding the focal subject and last at least 10 seconds (from the first behavioural indicator defining the context displayed by the conspecifics). The duration requirement was relaxed for three

of the contexts (play, aggression, and vigilance; see Table 3.2) which allowed us then to test for contextual consistency of the personality-linked behaviours of the chimpanzees. We classified the observations into eight naturally occurring and highly diverse contexts: feeding, play, grooming, resting, solitude, aggression, vigilance, and locomotion, (see Table 3.2 for details). Some contexts rarely occurred in our data, such as object manipulation or copulation, and were not assessed further. Only the first five contexts (feeding, play, grooming, resting, and solitude) were retained for the analyses because the last three were rarely observed in the data and represented less than 10% of the total duration of the complete dataset (see Table B.2). As previously done in personality studies (Kuhar et al., 2006; Pederson et al., 2005), grooming and play were merged to create an affiliative context as they together represented approximately 10% of the overall dataset. However, all eight contexts were considered when comparing the rating method (behaviour-descriptive item) with the naturalistic method on the behaviour level. For further information regarding the representativeness of the behavioural expression across individuals within each context and for each time period, see Chapter 3.

A naïve observer coded 20% of the complete dataset for reliability by measuring the behaviours independently from the contexts. The two coders reliably classified the same contexts within 3 seconds (i.e., margin of error), Cohen's Kappa ($\kappa = 0.79$). The two coders agreed on the type and number of behaviours, intraclass correlation coefficient (mean ICC(3,1) = 0.59 , SD = 0.24, $p<0.05$).

Rating approach

The personality traits of the subjects were rated between May and September 2017. The questionnaire was devised for this study based on a previous version developed by Uher

and colleagues (Uher & Asendorpf, 2008; Uher & Visalberghi, 2016). The questionnaire combined two assessment formats: behaviour-descriptive verbs and trait-adjectives. For the first format, the sentences were designed so that they represented a description of behaviours in specific situations. The second format used a single adjective that best described the personality trait assessed for. All items from both formats were randomly presented together regardless of the constructs that they measured. The rating was indicated using a 5-point Likert scale which ranged from 1 (almost never) to 5 (very often). In total, the questionnaire included 17 items (8 adjective items) to measure the four personality traits of interest (see Table C.2).

Hard copies of the questionnaires were translated in Bemba (i.e. Zambian local language) by locals and distributed to seven keepers who worked daily with the chimpanzees for 1 to 15 years (mean ± SD: 6.43 ± 5.00). Each keeper was asked to rate 10 chimpanzees and each chimpanzee was assessed by two to three keepers (2.92 ± 0.28). All keepers were questioned on the same day for the same individuals to prevent any discussions between them that could have possibly influenced their judgements. The keepers were asked to rate the same chimpanzees a second time after a break of five weeks. However, for the first session, due to management issues, one keeper rated the chimpanzees two weeks after the other keepers leading to a shorter break between the first and the second session.

Experimental approach

The experiments were conducted between May and September 2017 in the outdoor enclosure of the chimpanzees. An experiment was run only when the subjects had no other colony members within 15 meters. The tests were opportunistically performed at any time

between 6:30 am and 5:30 pm. Overall the data were collected evenly throughout the day with only three chimpanzees being tested only during the morning sessions. All subjects were tested twice for each experimental test, separating by at least a 5-week break.

Boldness was assessed with a rubber snake (SN), which could be associated with danger (see Figure 4.1a,b). The stimulus was hidden underneath a wire mesh and a pile of leaves, outside the enclosure within about 2 meters of the fence before any chimpanzees arrive in this area. During the test, the experimenter was standing 5 meters away from the stimulus while looking in the opposite direction of the stimulus, so the chimpanzee would not associate the experimenter with the snake. The experimenter slowly moved the snake by pulling a fishing line that was tied to the stimulus. The experiment lasted 1 minute, and 50 seconds and the chimpanzees were video-recorded. Each experiment was preceded by 1 minute of a baseline where the subject showed relaxed behaviours (e.g., sitting in a relaxed way, showing a relaxed face) while being within 5 meters of the fence, giving, then, the subject the opportunity to approach or step back from the stimulus moving along the fence.

Figure 4.1. Stimulus used to measure boldness: a) rubber snake hidden underneath the wire mesh and leaves, and b) rubber snake uncovered with the wire mesh aside.

Explorativeness was assessed with two novel objects (NO) (flashing dog toy hidden underneath a cardboard box; a set of tin cans piled up on a wooden stick; see Figure 4.2a,b). During the test, the experimenter, who was standing between 1 and 2 meters behind the outdoor fence, showed the object to the subject. For the first object, a cardboard box was put on the ground upside down with the dog toy hidden underneath; the experimenter was then pulling a string on each side of the box revealing the toy (see Figure 4.2a). For the second object, the tin cans were piled up on the wooden stick by the experimenter who poured water into each of them (see Figure 4.2b). The experiment lasted 2 minutes and the chimpanzees were video-recorded. Each experiment was preceded by 1 minute of a baseline where the subject showed relaxed behaviours while being within 5 meters of the fence. Each subject was tested once for each of the two objects. The order of presenting the objects was counterbalanced across subjects.

Figure 4.2. Objects used to measure exploratiVeness: a) flashing dog toy hidden underneath a cardboard box, and b) set of cans piled up on a wooden stick and water being poured in.

Six behaviours were measured depending on the experimental test (see Table 4.1). They were chosen based on previous experimental studies which targeted the personality traits of interest (Carter et al., 2012a; Kutsukake et al., 2012; Massen et al., 2013; Neumann et al., 2013). However, we made some adjustments; although the behaviour "approach" is often reported as a characteristic of exploratiVeness, we did not consider this behaviour for the NO test as most of our subjects were already sat/stood close enough (about 1 meter) to the fence before starting the experimental phase, preventing them to move forward. The behaviour "leave" was chosen over "move away" for the NO test as we believed that it was a stronger indicator of disinterest if the subject left the testing area. For each behaviour, we computed a frequency per minute of experimental test and were then standardized as z-score. To make sure that all behaviours reflected the same measurement, "move away" and "leave" were multiplied by -1, so that positive values would mean a higher level of boldness and exploratiVeness, respectively. Then, all behaviours that reflected the same personality trait were summed to create one score per personality trait per subject for each experimental test. Fifteen per cent of the videos were used for reliability purposes (mean $ICC(3,1) = 0.90$, $SD = 0.13$, $p<0.05$).

Table 4.1. Behaviours coded for the SN and NO tests and their definitions associated with the personality trait of interest.

Personality trait	Behaviour	Definition	Experimental test
Boldness	Approach	The subject walks towards the fence in the direction of the snake while moving within the 5-meter area	SN
	Move away	The subject walks away from the snake while moving within the 5-meter area	SN
Explorativeness	Gaze	The subject looks in the direction of the snake or the novel objects	SN, NO
	Straighten	The subject moves her/his upper body decisively in the direction of the snake or the novel objects while sitting or standing	SN, NO
	Leave	The subject moves outside of the 5-meter area	NO
Anxiety	Stress-related behaviours	The subject scratches her/him-self (gentle, rough), touches her/him-self (face, body), yawns, sways back and forth while sitting, shows bared-teeth display or emits distress call (hoo). The subject's hair rises up	SN, NO

Data Analysis

Before measuring the coherence between the different methods, we had to consider both contextual and temporal consistency (when applicable) to meet the criterion of personality assessment.

Reliability and consistency

To assess contextual and temporal consistency of the behaviours measured using the naturalistic and experimental methods, we used the Generalizability Theory (GT) (for more information on GT, see Chapter 2; Brennan, 2011; Shavelson & Webb, 1991; Shavelson, Webb, & Rowley, 1989). GT allows to identify multiple sources of systematic and unsystematic errors and estimate the variance components associated with each source of variation. We used a two-facet crossed design: subject by context by time, where subject

represents the object of measurement, and context (four levels), as well as time (two levels), represent the facets of generalization for the naturalistic method. Additionally, we used a one-facet crossed design – subject by time – for the naturalistic method when a behaviour was observed only in one context across two time periods, and for the experimental method. For our study, we conducted a Generalizability-study (G-study) on the different facets and their interactions; a coefficient of reliability was then obtained. Instead of having a strict cut-off to determine whether the coefficient was acceptable, we treated the coefficient per behaviour as a continuous variable to make use of most of the behaviours we measured in this study. Therefore, the behaviours showed either low, average or high consistency across context and time. For the naturalistic approach, we used a median split to determine which behaviours to retain for further analyses.

Regarding the rating data, out of 2380 questions in total, 18 were unanswered or indecisively answered (e.g., question marks or two different scores added) by four raters. It involved 9 items and 7 chimpanzees. For these missing values, we substituted the mean value for that item as previously done (Manson & Perry, 2000; Weiss et al., 2017). To measure the reliability of the ratings, the inter-rater reliability was measured using two intraclass correlation coefficient: ICC(3,1) and ICC(3,k). ICC(3,1) is used to estimate the reliability of ratings of any single rater which establishes how reliable would be a single rater if the ratings were only performed by them. ICC(3,k) estimates the reliability of the mean ratings across *k* raters (Shrout & Fleiss, 1979).

To assess the temporal consistency of the ratings, we used an approach that is commonly used in rating studies (Freeman et al., 2013; Weiss et al., 2015, 2017). We used the intraclass correlation coefficient (ICC) with a two-way mixed model – ICC(3,1), with the time period as the fixed variable and the individual as the random variable. ICC

measures the proportion of total variance in item that is due to differences between subjects while considering the within-subject variance (McGraw & Wong, 1996; Shrout & Fleiss, 1979). An item was considered consistent over time when the ICC value was significant ($p<0.05$).

Determining the personality structure

To determine the personality traits measured with the naturalistic and rating methods, we used principal component analyses (PCA). For the rating method, we ran two separate PCA for the trait-adjective and behaviour-descriptive verb items. To determine the number of components to extract, we used eigenvalue > 1 and the scree-plot. For the naturalistic method and ratings (adjective items), we used eigenvalue > 0.7, so we could extract four components that explained at least 80% of the variance. Additionally, we used an orthogonal rotation (Varimax). The behaviour/item loadings exceeding |0.5| were considered as salient (Budaev, 2010). In order to determine which type of rotation was the most appropriate, the analysis was repeated with an oblique rotation (direct Oblimin) which allowed the components to correlate. The correlations between the extracted components were low for the naturalistic approach (range: -0.03 – 0.34) and for the rating approach (range for adjective item: -0.34 – 0.08; range for behaviour-descriptor item: -0.24 – 0.10). Additionally, both Oblimin and Varimax rotations provided identical solutions regarding the behaviour and item (adjective and behaviour-descriptor verb) loadings. Therefore, we retained and interpreted the Varimax-rotated components. Subsequently, all behaviours and items (adjective and behaviour-descriptor verb) that loaded onto the same component (i.e., personality trait) were summed to create a single personality score for each subject per method (naturalistic and rating). The behaviours or items that loaded negatively onto the trait were multiplied by -1, so they would measure the same trait as the

other behaviours/items that loaded positively. The personality traits obtained with the rating method were then standardised as z-score allowing us direct comparisons with the other two methods.

To determine the personality trait structure of the individuals in the experimental method, we decided to use the conceptual approach (see section 2.3.2) as they were too few behaviours per experimental tests to carry out a PCA.

Coherence

To measure the coherence among the personality traits obtained with the naturalistic, experimental and rating methods, Spearman rank-order correlation coefficients (one-tailed) were used. We corrected the resulting p-values for multiple testing using Holm-Bonferroni correction (Holm, 1979).

All analyses were computed using SPSS Statistics 25 (IBM, Chicago, IL). GT analyses were run using the syntax developed by Mushquash and O'Connor (2006). In the syntax, all negative variances were set to zero by default. However, to obtain a more accurate idea of the reliability of the dataset, we ran additional GT analyses for the observational data (see Table B.5) and experimental data (see Table C.4) where all negative variances were kept in the results. RStudio Version 1.1.453 (RStudio, Boston, MA, USA) was used to assess the inter-rater reliability in the rating method using the "ICC" function ("psych" package; Revelle, 2019).

4.3 Results

Consistency assessment

Naturalistic approach

Three behaviours (Risky approach, Display, Throw) were discarded from the analyses as they were only expressed either in one time period or one context. G-study variance for the 19 behaviours is presented in Table 4.2 and included 22 chimpanzees.

Five behaviours were displayed only in one context across the two periods (Food begging, Food sharing, Risky action, Pilo-erection, Rock); none of them showed temporal consistency (G = 0; see Table 4.3), except for Food begging that showed high temporal consistency (G = 0.70; see Table 4.3). Among the 14 remaining behaviours, six of them (Escape, Playful contact, Chase, Gaze Approach, Gaze, Self-touch) had a variance of 0% suggesting a lack of inter-individual differences. The variance components attributable to subjects for the other eight behaviours ranged from 2.50% (Body contact) to 36.30% (Gaze Touch) suggesting inter-individual differences in their scores. The variance attributable to the main effect for measurement contexts, time and the interaction contexts by time accounted for less than 1% of the variance across the eight behaviours suggesting that the score for each behaviour within each subject was consistent across contexts and time periods when averaged over subjects and time or over subjects and contexts. The subject by context interaction accounted for 0% of variance for all behaviours but Self-scratch (4.30%), and the subject by time interaction accounted for 0% for all eight behaviours, indicating very little if not, variability across the individuals in their pattern of contextual and temporal consistency for the eight behaviours. Finally, the three-way subject by

context by time interaction (residual variance) accounted for a large percentage of variance; this variance reflects different behavioural score per subject across time and contexts as well as other sources of error not systematically included into the G-study.

Table 4.2. Estimated variance components for the G-study subject by context by time design and subject by time design per behaviour when considering the feeding (F), affiliative (GP), resting (R) and solitude (S) contexts.

Context	F-GP-R		F-GP-R		F-GP-R		F-GP		F-GP-R		F-GP-R		F-GP-R-S	
Behaviour	Body contact		Groom		Playful contact		Chase		Rough action		Gaze		Gaze Approach	
Effect	Var	%	Var	%	Var	%	Var	%	Var	%	Var	%	Var	%
Subject	0.03	2.50	0.25	18.00	0.00	0.00	0.00	0.00	0.29	21.70	0.00	0.00	0.00	0.00
Context	0.01	0.60	0.01	1.00	0.00	0.00	0.00	0.00	0.01	0.80	0.00	0.00	0.00	0.00
Time	0.00	0.00	0.00	0.10	0.00	0.00	0.00	0.00	0.00	0.20	0.00	0.00	0.00	0.00
Subject*Context	0.00	0.00	0.00	0.00	0.18	16.10	0.81	75.40	0.00	0.00	0.23	19.90	0.02	1.90
Subject*Time	0.11	9.30	0.00	0.00	0.05	4.10	0.00	0.00	0.00	0.00	0.19	16.70	0.16	14.90
Context*Time	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Subject*Context*Time	1.01	87.70	1.10	80.80	0.88	79.70	0.26	24.60	1.03	77.20	0.72	63.40	0.89	83.20
Context	F-GP-R-S		F-GP-R-S		F-GP		F-GP-R-S		F-GP-R-S		F-GP-R-S		GP-R-S	
Behaviour	Gaze Manipulate		Gaze Touch		Escape		Self-scratch		Self-touch		Vigilant		Yawn	
Effect	Var	%	Var	%	Var	%	Var	%	Var	%	Var	%	Var	%
Subject	0.28	28.00	0.40	36.30	0.00	0.00	0.16	15.30	0.00	0.00	0.30	27.80	0.34	28.70
Context	0.00	0.00	0.00	0.30	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.40
Time	0.00	0.00	0.00	0.10	0.00	0.00	0.00	0.20	0.00	0.00	0.00	0.40	0.00	0.20
Subject*Context	0.03	3.10	0.00	0.00	0.01	0.50	0.05	4.30	0.10	10.20	0.00	0.00	0.00	0.00
Subject*Time	0.10	9.60	0.00	0.00	0.19	17.90	0.00	0.00	0.21	21.00	0.00	0.00	0.00	0.00
Context*Time	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Subject*Context*Time	0.59	59.30	0.70	63.20	0.88	81.60	0.83	80.20	0.69	68.90	0.79	71.80	0.84	70.60

Comparison of methods to assess personality in sanctuary chimpanzees

Context	F		F		F		GP		S	
Behaviour	Food begging		Food sharing		Risky action		Pilo-erection		Rock	
Effect	Var	%	Var	%	Var	%	Var	%	Var	%
Subject	0.55	54.20	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Time	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Subject*Time	0.46	45.80	1.10	100.00	1.05	100.00	1.05	100.00	1.07	100.00

Var: Variance estimate was calculated using the mean squares; High values suggest that there is a large effect of the variable (see section 2.2.3, Table 2.1). %: Percentage of the total variance for each variance estimate was calculated by dividing each variance estimate by the total variance.

Among the eight behaviours that showed inter-individual variance across both contexts and time, the G-coefficients varied from 0.34 to 0.82 indicating low to high-reliability scores overall (see Table 4.3). Following Figueredo and colleagues (1995), we constructed the G-coefficients for context and time separately for each of the eight behaviours (see Table 4.3). For context, the G-coefficients ranged from 0.78 to 1, indicating high reliability across feeding, affiliative, resting and solitude contexts when averaged over subjects and time. Concerning time, the G-coefficients varied from 0.21 to 1, indicating low to high reliability over time when averaged over subjects and contexts.

Table 4.3. Overview of the G-coefficients per behaviour for subject by context by time design and for subject by time design when considering the feeding (F), affiliative (GP), resting (R) and solitude (S) contexts. The table depicts the G-coefficients for both context and time separately.

Context	Behaviour	G-coefficient	G-coefficient for context	G-coefficient for time
F-GP-R	Body contact	0.34	1.00	0.21
F-GP-R	Groom	0.57	1.00	1.00
F-GP-R	Playful contact	0.00	-	-
F-GP	Chase	0.00	-	-
F-GP-R	Rough action	0.63	1.00	1.00
F-GP-R	Gaze	0.00	-	-
F-GP-R-S	Gaze + Approach	0.00	-	-
F-GP-R-S	Gaze + Manipulate	0.68	0.90	0.74
F-GP-R-S	Gaze + Touch	0.82	1.00	1.00
F-GP	Escape	0.00	-	-
F-GP-R-S	Self-scratch	0.58	0.78	1.00
F-GP-R-S	Self-touch	0.00	-	-
F-GP-R-S	Vigilant	0.76	1.00	1.00
GP-R-S	Yawn	0.71	1.00	1.00
F	Food begging	0.70	-	-
F	Food sharing	0.00	-	-
F	Risky action	0.00	-	-
GP	Pilo-erection	0.00	-	-
S	Rock	0.00	-	-

The additional GT analysis which replaced the variances of 0 with the original negative values revealed a similar pattern for the different behaviours, although the values were larger overall (see Table B.5). The variance components attributable to subjects that were 0% for Food sharing, Risky action, Pilo-erection, Rock, Escape, Playful contact, Chase, Gaze, Gaze Approach and Self-touch had negative values which could suggest a problem in the dataset; i.e., low amount of data point per subject which may lead to very little variation between individuals. The variance attributable to the main effect for measurement contexts and time had a small negative value (less than - 0.10) for some of

the behaviours (Body contact, Gaze Manipulate, and Self-scratch). As recommended in the literature (see Chapter 2 for details), these values can be set to zero, suggesting consistency over time when averaged over subjects and contexts. With regard to the interactions between the different facets (context and time), for some of the behaviours (Body contact, Groom, Rough action, and Yawning), the negative values had a bigger magnitude than for the main effect measurements, ranging from - 0.311 to - 0.114. Finally, the variance for the three-way interactions is relatively high for some of the behavioural actions as a result of the presence of negative values for the main effect and interaction measurements. All behaviours that were originally selected for the PCA would also be selected if using this second set of GT analysis.

As mentioned in the previous analysis, the interpretation of the consistency for some of the behaviours has to be considered carefully here; however, a bigger dataset would likely provide stronger patterns of consistency for most of the behaviours.

Rating approach

The inter-rater reliability of the behaviour-descriptive item "If there is a new object (like when the water or juice fountain was built) in the enclosure, __touches it" was equal to 0 in both sessions; hence, this item was not retained for further analyses. The ICC(3,1) estimates of the adjective items ranged from 0.03 to 0.22, and from 0 to 0.39 for session 1 and 2, respectively. The ICC(3,1) estimates of the behaviour-descriptive items ranged from 0 to .35 and from 0 to 0.58 for session 1 and 2, respectively (see Table 4.4). Overall, the inter-rater reliability scores of single ratings for both adjective and behaviour-descriptive items are fairly low, although previous nonhuman primate research reported similar values

(Weiss et al., 2015, 2017). Only the items that were reliable at least in one of the sessions were retained for further analyses; in total 16 items (8 adjective items) were retained.

Table 4.4. Inter-rater reliability scores for the adjective and behaviour-descriptive items. The table provides both ICC(3,1) and ICC(3,k) for each session. Items retained for further analyses are in bold.

Format	Item	Session 1		Session 2	
		ICC(3,1)	ICC(3,*k*)	ICC(3,1)	ICC(3,*k*)
Adjective	**_is friendly to other chimps.**	0.10	0.44	0.29	0.74
	_is social.	0.03	0.20	0.15	0.55
	_is bold.	0.20	0.64	0.39	0.82
	_is cheeky.	0.22	0.66	0.32	0.76
	_is curious.	0.08	0.37	0.00	0.00
	_is eager to learn.	0.17	0.59	0.07	0.36
	_is anxious.	0.32	0.77	0.27	0.72
	_is easily alarmed.	0.04	0.23	0.00	0.00
Behaviour-descriptive verb	**When relaxing, _is nearby other chimps.**	0.23	0.67	0.08	0.36
	During feeding, _sits nearby other chimps.	0.00	0.00	0.06	0.30
	When playing with stronger chimps, _plays rough and chases them.	0.35	0.79	0.58	0.91
	When eating next to stronger chimps, _takes food from them.	0.00	0.00	0.52	0.88
	If there is a new object (like when the water or juice fountain was built) in the enclosure, _touches it.	0.00	0.00	0.00	0.00
	If there is a music box in the enclosure, _approaches and touches it.	0.02	0.11	0.00	0.00
	If there is a flashing toy in the enclosure, _approaches and touches it.	0.04	0.24	0.00	0.00
	When waiting for food, _scratches him/herself or walks back-and-forth.	0.17	0.59	0.05	0.25
	When others fight, _screams or runs away.	0.12	0.48	0.12	0.49

Using the mean ratings of the two to three raters per chimpanzee, we measured the temporal consistency of the 16 reliable items. The mean (± SD) temporal consistency of the 16 items was 0.37 (± 0.25) (range: 0.00 – 0.74), $p<0.05$. Only 10 of them (five adjective items) showed temporal consistency (see Table 4.5).

Table 4.5. Overview of the temporal stability for each adjective and behaviour-descriptive item associated with the personality trait of interest. Items that showed temporal consistency are in bold.

Format	Item	ICC(3,1)	95% Confidence Interval		F value	Sig
			lower bound	upper bound		
Adjective	**_is friendly to other chimps.**	0.41	0.02	0.70	2.40	**0.020**
	_is social.	0.62	0.30	0.82	4.30	**< 0.001**
	_is bold.	0.67	0.38	0.84	5.10	**< 0.001**
	_is cheeky.	0.57	0.23	0.79	3.70	**0.001**
	_is curious.	0.00	-0.40	0.40	1.00	0.500
	_is eager to learn.	0.00	-0.40	0.40	1.00	0.500
	_is anxious.	0.32	-0.08	0.64	2.00	*0.057*
	_is easily alarmed.	0.45	0.06	0.72	2.60	**0.013**
Behaviour-descriptive	When relaxing, _is nearby other chimps.	0.23	-0.18	0.58	1.60	0.130
	During feeding, _sits nearby other chimps.	0.34	-0.07	0.65	2.00	**0.048**
	When playing with stronger chimps, _plays rough and chases them.	0.74	0.48	0.88	6.60	**< 0.001**
	When eating next to stronger chimps, _takes food from them.	0.66	0.36	0.84	4.90	**< 0.001**
	If there is a music box in the enclosure, _approaches and touches it.	0.00	-0.40	0.40	1.00	0.500
	If there is a flashing toy in the enclosure, _approaches and touches it.	0.15	-0.27	0.51	1.34	0.240
	When waiting for food, _scratches him/herself or walks back-and-forth.	0.35	-0.06	0.65	2.10	**0.044**
	When others fight, _screams or runs away.	0.44	0.05	0.71	2.60	**0.014**

For the following analyses, in order to make use of all available data, a mean value aggregating the two sessions was obtained for the ten items that were consistent over time. For the other six items, we retained the value from session 1.

Experimental approach

We tested for temporal consistency for both SN and NO tests. Although the SN and NO tests were designed to target boldness and explorativeness, respectively, based on the behaviours displayed by the subjects, we also considered explorativeness and anxiety for the SN test, and anxiety for the NO test (see Table C.3; Figure 4.3).

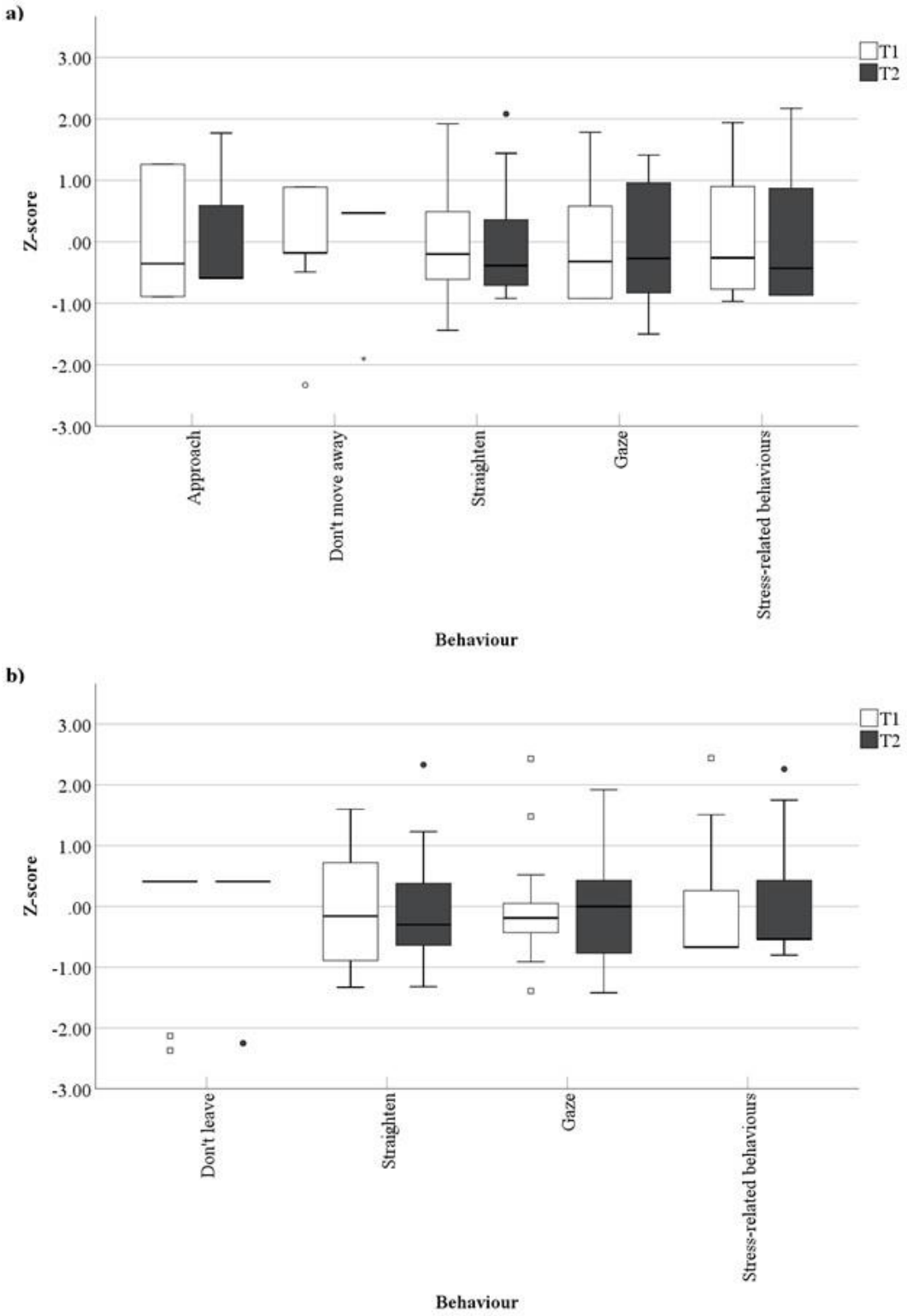


Figure 4.3. Visual representation of the behaviours displayed across subjects per time period during the a) SN and b) NO tests. Data are z-scored within each time period. T1 and T2 represents the time period 1 and 2, respectively. The thick horizontal lines indicate medians; the vertical length of the boxes corresponds to interquartile range; the thin short horizontal lines indicate the minimum and maximum values.

For the SN test, only anxiety showed a high level of temporal consistency (G-coefficient = 0.67). Boldness did not show any variance across the subjects indicating a lack of inter-individual differences, and exploratiVeness only reached 3.2% suggesting very little inter-individual differences. Their G-coefficients were very low suggesting a lack of consistency over time.

For the NO test, the variance components attributable to subjects were high for both exploratiVeness and anxiety, suggesting inter-individual differences in scores. Both traits showed high level of temporal consistency (G-coefficients > 0.80; see Table 4.6).

Table 4.6. Estimated variance components for the G-study subject by time design as well as the G-coefficients per personality trait for both SN and NO experimental tests.

		Anxiety		Boldness		Explorativeness	
Experimental test	Effect	Var	%	Var	%	Var	%
SN	Subject	0.50	50.10	0.00	0.00	0.06	3.20
	Time	0.00	0.00	0.00	0.00	0.00	0.00
	Subject*Time	0.50	49.90	2.54	1.00	1.81	96.80
	G-coefficient	0.67		0.00		0.06	
NO	Subject	0.70	70.20	-	-	2.59	70.40
	Time	0.00	0.00	-	-	0.00	0.00
	Subject*Time	0.30	29.80	-	-	1.09	29.60
	G-coefficient	0.83		-		0.83	

As previously found for the naturalistic approach, the additional GT analysis revealed a similar pattern for the different personality traits in each experimental test, although the values were larger overall (see Table C.4). The variance component attributable to subjects that was 0% for Boldness had a negative value which could suggest a problem in the dataset; i.e., little variation between individuals for one of the time period (see Figure 4.3). The variance attributable to the main effect for measurement time had a small negative value (less than - 0.10) for Anxiety in SN and for both Anxiety and

Explorativeness in NO. As recommended in the literature (see Chapter 2 for details), these values can be set to zero, suggesting consistency over time when averaged over subjects. With regard to Explorativeness in SN, the negative value for the time measurement had a larger magnitude (- 0.181), suggesting possibly a problem in the data (e.g., lack of consistent behavioural responses). This lack of consistency is not entirely surprising given the subjects saw the same snake for a second time which possibly resulted in the loss of the novelty aspect. The G-coefficients for each personality trait were the same values as the original analysis.

Data reduction

Naturalistic approach

We used a split-median to determine which behaviours were retained for further analyses. Based on the G-study results, all behaviours that were displayed in more than one context and obtained a coefficient above the median (M = 0.46) were averaged over time and across contexts.

A total of seven behaviours were retained for the PCA. The Kaiser–Meyer–Olkin (KMO = 0.545) and Bartlett's Test of Sphericity ($p < 0.001$) indicated the suitability of the dataset. We extracted four components which explained 86.7% of the variance (see Table 4.7). The first component explained 30.3% of the variance and had positive loadings of two behaviours that related to explorative actions; this component was labelled "exploratiVeness". The second component explained 21.9% of the variance and had positive loadings of two behaviours; one behaviour was related to social actions and the other one was associated with anxiety-related actions. This component was therefore

labelled "sociability". The last two components accounted for 19.9% and 14.6% of the variance, respectively. The first one had positive loadings of two behaviours that were related to anxious actions; we labelled this component "anxiety". Only one behaviour positively loaded onto the last component and was associated with bold actions; we labelled the component "boldness".

Table 4.7. Behaviour loadings after Varimax rotation when considering the feeding, affiliative, resting and solitude context.

Behaviour	Exploratiyeness	Sociability	Anxiety	Boldness
Groom		0.768		
Rough action				0.989
Gaze Manipulate	0.957			
Gaze Touch	0.968			
Self-scratch			0.908	
Vigilant			0.679	
Yawn		0.906		

Only loadings above |0.5| are reported.

Rating approach

For the items that were consistent over time, we obtained a mean value averaging the two sessions. For the items that were not consistent, we retained the value from session 1 in order to make use of all available data. We ran two separate PCA for the adjective and behaviour-descriptive verb items in order to determine how these two rating formats correlate with the other two methods (naturalistic and experimental observations).

Following the ICC analysis, 8 adjective items were retained for the PCA. The Kaiser–Meyer–Olkin (KMO = 0.652) and Bartlett's Test of Sphericity ($p < 0.001$)

indicated the suitability of the data. We extracted four components which explained 88.8% of the variance (see Table 4.8). The first component explained 24.2% of the variance and had positive loadings of adjective items that related to bold description; this component was labelled "boldness". The second component explained 24% of the variance and had positive loadings of adjective items that were related to both explorative and social description; this component was therefore labelled "explorativeness-sociability". The last two components accounted for 19.3% and 17.3% of the variance, respectively. The first one had a positive loading of "anxious" and a negative loading of "easily alarmed". We initially thought that these two items measured the same personality trait and would, therefore, load in the same direction onto the component. We decided to label this component "anxiety" based on the item "anxious" which represents a clear reflection of this trait. Only one item positively loaded onto the last component which referred to a social adjective; we labelled the component "sociability".

Table 4.8. Adjective item loadings after Varimax rotation.

Adjective item	Boldness	Explorativeness-Sociability	Anxiety	Sociability
Social		0.714		
Is friendly to other chimps				0.943
Bold	0.893			
Cheeky	0.893			
Curious		0.707	(0.511)	
Eager to learn		0.849		
Anxious			0.731	
Easily alarmed			-0.741	

Only loadings above |0.5| are reported. The item with the value in brackets was not associated with this component as it loaded higher on another one.

Following the ICC analysis, 8 behaviour-descriptor verb items were retained for the PCA. The Kaiser–Meyer–Olkin (KMO = 0.601) and Bartlett's Test of Sphericity (p =

0.001) indicated the suitability of the data. We extracted three components which explained 69.1% of the variance (see Table 4.9). The first component explained 29% of the variance and had positive loadings of descriptive items that related to bold and social behaviours. The social-related descriptive item positively loaded onto the second component as well; however, as its loading was higher on the first component, we retained this association. This component was labelled "boldness-sociability". The second component explained 24.1% of the variance and had positive loadings of "When relaxing […] chimps" as well as "When others fight […] away", and a negative loading of "When waiting for food […] back-and-forth". We expected the last two items to measure anxiety, however, they did not load in the same direction suggesting that they do not measure the same aspect of the trait. As this component had positive loadings of two items that relate to social behaviour (although one was not retained) and had a stronger loading of "When waiting for food […] back-and-forth" compared to "When others fight […] away", we decided to label this component "sociability-no anxious". The last component accounted for 16% of the variance and had positive loadings of descriptive items that referred to explorative behaviours; hence, we labelled this component "explorativeness".

Table 4.9. Behaviour-descriptor verb item loadings after Varimax rotation.

Behaviour-descriptor verb items	Boldness-Sociability	Sociability-No anxious	Exploratiiveness
When relaxing, _ is nearby other chimps		0.675	
During feeding, _sits nearby other chimps	0.636	(0.535)	
When playing with stronger chimps, _ plays rough and chases them	0.940		
When eating next to stronger chimps, _ takes food from them	0.938		
If there is a music box in the enclosure, _approaches and touches it			0.672
If there is a flashing toy in the enclosure, _approaches and touches it			0.829
When waiting for food, _ scratches him/herself or walks back-and-forth		-0.829	
When others fight, _screams or runs away		0.650	

Only loadings above |0.5| are reported. The value in brackets was not retained as belonging to this component as it loaded higher onto a different one.

Method coherence

Trait level analysis

After correction for multiple tests, there were two significant positive correlations between the two rating formats for boldness, and between the experimental and rating (adjective item) approaches for exploratiiveness (see Table 4.10). Additionally, there was a positive trend between the naturalistic and the adjective rating approaches for sociability and two negative trends between the naturalistic and adjective rating approaches as well as between the naturalistic and the SN test for anxiety.

Table 4.10. Spearman rank coefficients between the personality traits measured within each method. Value in bold represents significant correlation after Holm-Bonferroni correction.

Personality trait	R_A–R_B	Nat–R_A	Nat–R_B	Nat–E_{SN}	Nat–E_{NO}	R_A–E_{SN}	R_A–E_{NO}	R_B–E_{SN}	R_B–E_{NO}
Sociability	.16	.42*	0.32 (0.16)#						
Boldness	**.82****	0.28	0.13						
Explorativeness	.23	-0.03	-0.08		-0.48		**.67****		-0.09
Anxiety	.31	-.44*	-0.06	-.64*	-0.14	-0.130	0.27	-0.04	0.11

Nat: naturalistic approach; R_A: trait-adjective items; R_B: behaviour-descriptive verb item; E_{SN}: snake test; E_{NO}: novel object test. #The value in brackets represents the coefficient when correlating "sociability" measured with the naturalistic approach and "sociability-unanxious" measured with the behaviour-descriptive verb rating approach. *p<0.05, **p<0.01. N = 24 for correlations between trait-adjective and behaviour-descriptive verb items; N = 22 for correlations between naturalistic and rating approaches; N = 9 and N = 12 for correlations between naturalistic approach and SN test and NO test, respectively; N = 13 and N = 10 for correlations between rating approach and NO test and SN test, respectively.

Behaviour level analysis

To obtain a more detailed picture with regards the coherence between the naturalistic and rating approach, we compared both methods on the behaviour level. Hence, for the naturalistic approach, we selected behaviours displayed in a specific context to match with the behaviour-descriptive item used in the rating approach. For exploratiVeness, we only considered explorative behaviours occurring in neutral contexts (i.e., resting and solitude) since the two exploratiVeness-related items did not have a specific situation associated with. For both methods, each behaviour and items values were standardized using z-score to allow the comparison between the two methods.

The analyses revealed two significant correlations between the items and the matching context/behaviour (see Table 4.11). First, the chimpanzees who were rated high on the item "When relaxing […] chimps" were the ones who displayed less body contacts while resting when measured with the naturalistic approach. Second, the chimpanzees who were rated high on the item "When playing […] chases them" were also the ones showing rough and chasing behaviour when displayed in an affiliative context. Additionally, we had three positive trends for sociability and anxiety-related behaviours/items suggesting possible coherence between the two methods.

Table 4.11. Spearman rank correlation coefficients between the behaviour displayed within a specific context and the matching behaviour-descriptive item. Values in bold represent significant correlations after Holm-Bonferroni correction (when applicable).

		Rating method (Behaviour-descriptive items)						
	Context and behaviour	When relaxing, _ is nearby other chimps	During feeding, _sits nearby other chimps	When playing with stronger chimps, _ plays rough and chases them	When eating next to stronger chimps, _takes food from them	If there is a music box in the enclosure, _approaches and touches it	If there is a flashing toy in the enclosure, _approaches and touches it	When waiting for food, _ scratches him/herself or walks back-and-forth
Naturalistic method	Resting_Body contact	**-0.36***						
	Resting_Groom	0.41*						
	Feeding_ Body contact		0.37*					
	Feeding_Groom		-0.07					
	Affiliative_Chase/Roug action			**0.58***				
	Feeding_Risky action				-0.17			
	Neutral_Explorative behaviour					0.01	0.24	
	Vigilance_Self-scratch							0.14
	Vigilance_Vigilant							0.43*

*p<0.05

4.4 Discussion

This study compared three methods of personality assessment – naturalistic, rating and experimental – measuring traits of sociability, boldness, explorativeness and anxiety of sanctuary chimpanzees. We aimed to determine whether each of these three methods measured the same facet of expression of the personality traits targeted. We also compared the methods in specific items (behaviour-descriptive verb) to explore potentially different facets of the same traits. On the trait level, there were significant positive correlations between the two formats of ratings for boldness, and between the rating (adjective) and experimental methods for explorativeness suggesting that the methods targeted the same personality trait. On the behaviour level, the rating and naturalistic methods showed a significant positive correlation for boldness, which implies that both methods targeted the same facet of expression of boldness. In contrast, we found a negative correlation for sociability which suggests that the two methods did not target the same facet of expression.

In this study, when comparing the rating (adjective) with the naturalistic approach, we found no significant relations but a positive trend in sociability supporting previous research (Freeman et al., 2013; Tkaczynski et al., 2018; Uher & Visalberghi, 2016; Uher, Werner, et al., 2013). In contrast, we found a negative trend in anxiety. It is important to note that the keepers mostly observe the chimpanzees when they provide the food during the afternoon feeding. Previous research reported that a pre-feeding period can lead to stress-related behaviours and aggression due to the anticipation for food (Palagi, Cordoni, & Borgognini Tarli, 2004; Waitt & Buchanan-Smith, 2001; Yamanashi, Nogami, Teramoto, Morimura, & Hirata, 2018). Consequently, it is possible that the keepers may have been influenced while estimating the chimpanzees' anxiety level for some of the

individuals. Regarding exploratiVeness, past research reported positive correlations when comparing these two methods (Uher, Werner, et al., 2013); here, this lack of coherence may indicate that both methods target a different facet of the trait. In the naturalistic approach, exploratiVeness was measured towards objects, food as well as conspecifics covering different aspects of being explorative. It is possible that keepers may associate being "curious" or "eager to learn" to interactions with objects or food only. Indeed, here, we found a positive correlation between the rating (adjective) and experimental methods for exploratiVeness. The interpretation of a rater seems to play an important part when assessing the personality of individuals as previously reported for capuchin monkeys (Uher & Visalberghi, 2016). Our findings suggest that considering distinctive contexts plays an important role when it comes to comparing different methods for personality research.

Our analysis on the behaviour level comparing the naturalistic and rating method (behaviour-descriptiVe verb) revealed interesting patterns. As predicted, the two methods showed significant coherence in their assessment for boldness, and strong tendencies for sociability (when considering the behaviour "groom") and anxiety. However, we found a significant negative correlation for sociability when considering body contact as a key behaviour to determine this trait. Because grooming is a more noticeable behaviour to observe than body contact, it is possible that the keepers may associate a relaxing context with grooming events which could explain the positive relationship observed between resting context/groom and the matching item. In contrast, it is less likely to observe grooming behaviour during feeding time, but, instead, body contacts are probably more likely to occur than grooming behaviour due to the nature of the context. This could explain the positive relationship when comparing feeding-body contact with the matching item. Although previous research did not compare the two methods with such a detailed

approach, similar positive relations were observed for the same personality traits in chimpanzees (Freeman et al., 2013) and crab-eating macaques (Uher, Werner, et al., 2013). Breaking down the two methods allowed us to demonstrate that both naturalistic and rating (behaviour-descriptive items) methods can target the same facet of expression of a personality trait.

With regard to the comparison between the experimental observations and ratings, we expected to see higher coherence when compared to the behaviour-descriptive verb items due to their close association to a specific situation (e.g., Massen et al., 2013; Uher & Asendorpf, 2008). However, here, our analysis revealed a different pattern in the data. A significant positive relation was found with ratings (adjective) for exploratiVeness. Uher and colleagues previously found coherence between experimental approach and both adjective and behavioural-descriptive item rating approach in capuchins, macaques and great apes (Uher & Asendorpf, 2008; Uher & Visalberghi, 2016; Uher, Werner, et al., 2013). The fact that the exploratiVeness score measured with the adjective rating method matched with the score measured with the NO test suggests that both methods assessed the same facet of the trait targeted. This finding could imply that people may perceive exploratiVeness as a reflection of showing interest towards an object. To help determine how comparable different methods are when assessing animal personality or better estimate the diversity of trait expression, future studies should include more in-depth in their analyses.

The analyses comparing the behavioural observations in the natural and experimental setting for exploratiVeness and anxiety revealed a negative trend for anxiety when compared in the natural setting and the SN test. This finding suggests perhaps that

they measured different facets of expression of the same trait. For the naturalistic approach, no threatening contexts were included when determining the personality structure. So, it is possible that anxiety-related behaviours measured experimentally in a threatening situation (i.e., snake test) elicited different patterns of stress behaviours compared to any other contexts (feeding, affiliative, resting or solitude). Additionally, the use of a snake model to capture boldness-related responses might have been too inconsistent with the behaviours measured in the naturalistic method, especially because the observational data included a broader range of behaviours displayed in various contexts than for the snake model. A lack of coherence between experimental and naturalistic methods for different traits (sociability, boldness, anxiety, connectedness, aggressiveness) was previously reported in crested macaques (*Macaca nigra*) (Neumann et al., 2013); however, the authors did not compare the exact same personality traits across methods as it was not their main objective. The naturalistic approach provided a much more detailed picture of the two personality traits compared to the experimental approach that was associated with a specific situation. For instance, in the NO test, the subjects were not able to touch or manipulate the objects for technical reasons, giving, thus, a different perspective in their behavioural reactions towards the object. It is possible that when comparing behavioural measures in naturalistic approach to experimental or rating approaches, the lack of coherence found in previous research (e.g., Neumann et al., 2013; Tkaczynski et al., 2018) may have been due to the lack of focus on the contexts in the natural settings.

Concerning the correlations between the two rating formats, we found only one significant positive correlation for boldness. This finding is in line with previous research in crab-eating macaques (Uher, Werner, et al., 2013). The lack of significant relationships

between the two formats for the other three personality traits is not a complete surprise as previous research in capuchins (Uher & Visalberghi, 2016) and great ape species (Uher & Asendorpf, 2008) reported similar results for sociability and anxiety. However, these previous studies reported significant positive relationships between the two formats of rating for exploratiVeness (Uher & Asendorpf, 2008; Uher & Visalberghi, 2016; Uher, Werner, et al., 2013). The lack of coherence may suggest that they do not target the same facet of expression of the traits, thus they are not interchangeable.

In the current study, within each method, all personality traits were reliably measured – i.e., both naturalistic and experimental showed contextual or temporal consistency, and the keepers agreed in their assessment for 16 out of 17 items. Our study revealed a certain variability in terms of coherence across the methods, suggesting that each of these methods may distinctively target different facets of expression of the same personality trait. Alternatively, the lack of coherence between the methods may also suggest that they do not measure the same personality trait. It is important to note that in our study, only a few behaviours met the key criteria to be retained to create the personality traits when using the naturalistic approach, which is possibly due to the few data points per subject. It would be interesting in future studies to use perhaps a more conceptual approach where more diverse behaviours are used to generate the different personality traits in comparison with the approach used here. By doing so, this conceptual approach may provide different insight regarding the analyses comparing the different methods (Uher & Asendorpf, 2008; Uher & Visalberghi, 2016). Additionally, conducting repeated bivariate correlations to assess the coherence between the methods might not have been the most appropriate statistical test to determine whether the personality traits are consistently measured across all methods. Instead, future studies could perhaps use a GT

analysis and test whether the personality trait scores generalize across different methods of assessment.

Previous research reported that some traits may be divided into multiple sub-traits (Carter et al., 2012a; Koski, 2011a; Tkaczynski et al., 2018). These findings could explain why we found a lack of coherence between the methods for some of the personality traits. For instance, it has been suggested that sociability could be characterised by multiple sub-traits such as tactility, equitability, positive affect and a more general sociability trait (Koski, 2011a; Tkaczynski et al., 2018) which implies, to some degree, the importance of the context in which the personality trait is expressed. Indeed, for instance, tactility was associated with contexts such as grooming (Tkaczynski et al., 2018) whereas positive affect related to a play context (Koski 2011b). Therefore, it is essential that future research consider the importance of the contexts in their analysis, so they could provide a detailed picture regarding the comparisons for the different methods of assessment.

To conclude, the three approaches of this study seem to reliably measure sociability, boldness, exploratiVeness and anxiety of sanctuary chimpanzees. Our study revealed that for some personality traits, these three methods show coherence in their assessment, especially when considering the context and behaviour of the naturalistic approach. Unlike previous research which associated a lack of coherence between methods with a lack of equivalence in the methodology, we suggest, instead, that each approach may measure different facets of expression of the same personality trait, and some of these facets may be targeted by the different methods. Individuals expressed their personality traits through different behavioural actions in diverse contexts throughout their lives. Thus,

it is important to develop methods that can capture this diversity of personality expression in order to better understand the complexity of an individual's personality.

5. Personality measured in daily contexts of rehabilitant Bornean orangutan (Pongo pygmaeus) using naturalistic observations

Abstract

Within species of nonhuman primate, individuals may show differences from one another in the expression of behaviour-based traits; these inter-individual differences have been found to be consistent across time and contexts in past research. In many species, including orangutans, inter-individual differences have been examined using ratings and experimental observations. In the current study, we used naturalistic observations of rehabilitant orangutans (n=20), to test for the consistency of 17 behaviours linked to the personality traits of sociability, boldness, exploratiVeness and anxiety that occurred across two to seven distinct daily contexts and two time periods. Eight behaviours showed consistency across contexts and time; these behaviours were then subjected to a principal component analysis which revealed two components: anxiety and a combination of sociability/boldness/exploratiVeness. This study, thus, provides the first evidence of contextual and temporal consistency in orangutans' behaviours in naturalistic observations. On a preliminary level, we also examined to what extent age and socioecological factors can influence the contextual consistency. Our study revealed differences in the pattern of contextual consistency across subjects. These preliminary findings suggest that developmental and/or socioecological factors may influence the expression of the personality-linked behaviours of these orangutans. Using a naturalistic approach that

captures different trait-based behaviours occurring across multiple contexts and time can help improve our understanding of how nonhuman primates express their personality in their day-to-day lives and possibly develop further our insight about the complexity of measuring personality.

5.1 Introduction

Previous studies on animal personality showed that individuals may differ from one another in the way they express their behaviour-based traits, which constitute their personality, across multiple situations and different periods of time (Gosling, 2001, 2008). These differences were reported to be consistent over time and across different contexts (Réale et al., 2007; Sih, Bell, & Johnson, 2004; Sih, Bell, Johnson, et al., 2004). If we want to understand the full expression of the personality traits of a species, it is important to test for both temporal and contextual consistency. Such consistencies have been demonstrated for a wide range of taxa, arguably ranging from invertebrates (e.g., Chapman, Thain, Coughlin, & Hughes, 2011; Hewes & Chaves-Campos, 2018) to vertebrates (Réale, Gallant, Leblanc, & Festa-Bianchet, 2000b; Schuster, Carl, & Foerster, 2017), with nonhuman primates receiving special attention [prosimians: (Dammhahn, 2012); New World monkeys: (Šlipogor et al., 2016); Old World monkeys: (Carter et al., 2012b); great apes: (Weiss et al., 2012)]. Conducting research on different personality traits in various nonhuman primates, and particularly in all four great ape species (orangutans, gorillas, bonobos and chimpanzees), could provide an important contribution to better understand the evolutionary history of human personality (Buss, 1988; Nettle, 2006)

The temporal consistency of personality traits has been demonstrated in various primate species for both short (e.g., a few weeks apart: Tomassetti et al., 2019; Uher, Asendorpf, & Call, 2008) and long-time periods (e.g., years apart: Koski, 2011; Suomi, Novak, & Well, 1996). In contrast, contextual consistency tended to be overlooked in research on nonhuman primates, even though personality may be expressed across multiple daily contexts (Chapter 3). It is also important to note that when contexts were examined, the literature on nonhuman primates tended to focus mostly on very specific situations in an experimental setting (Carter et al., 2012b; Dammhahn & Almeling, 2012; Massen et al., 2013; Šlipogor et al., 2016; Uher et al., 2008; Uher, Werner, et al., 2013). For instance, predator-like situations (e.g., snake vs. leopard) and novel objects (e.g., small vs. big) were often used to test for contextual consistency of boldness and explorativeness, respectively (Massen et al., 2013; Šlipogor et al., 2016; Uher et al., 2008). As personality traits can be expressed in multiple ways and various situations, an important current step to take in nonhuman primate personality research is to target a wider range of behaviours as well as contexts in order to better understand the structure of personality. Naturalistic observations allow us to observe individuals' behaviours in the course of their lives where they can interact with both their physical and social environments, providing us with real-life data.

Because individuals can express their personality traits through diverse behaviours – e.g., sociability can be measured by the ability of an individual to share food (Silk, Brosnan, Henrich, Lambeth, & Shapiro, 2013) or to sit in proximity of others (Massen & Koski, 2014) – in multiple distinct contexts and throughout their lives, personality remains a complex concept to understand fully and study, in terms of behaviours, context and time. Therefore, if we want to grasp the diversity of expression of different traits, we need to measure the behaviours of animals on multiple occasions and in multiple contexts. The naturalistic

approach is likely to contribute to a better understanding of the personality structure and complexity of the individuals.

To our knowledge, only 10 studies on nonhuman primates examined personality consistency using exclusively naturalistic observations, which was carried out on altogether six primate species (Bardi et al., 2001; Brent et al., 2014; Byrne & Suomi, 1995; Koski, 2011b; Locke, Locke, Morgan, & Zimmermann, 1964; Maestripieri, 2000; Reite & Short, 1980; Seyfarth et al., 2012; Suomi et al., 1996; von Borell et al., 2016), but none of them examined contextual consistency. To date, only one study, carried out on sanctuary chimpanzees (*Pan troglodytes*), examined and provided evidence for contextual and temporal consistency (Chapter 3). In this study, consistency was found across multiple distinct contexts that did not always share commonalities on the functional, affective, arousal or social level (e.g., resting, feeding and solitude contexts for anxiety) highlighting the diversity of expression of a personality trait. There is a clear need to further test for personality consistency, particularly contextual consistency, in nonhuman primate species by using naturalistic observations. Examining orangutan personality with such an approach with high ecological validity would help better understand how these great apes express their personality on a day-to-day basis.

Whereas much personality research has focused on chimpanzees (Dutton, 2008; King et al., 2005; Koski, 2011b; Kutsukake et al., 2012; Massen et al., 2013; Uher et al., 2008; Weiss et al., 2007, 2017), the great apes that are together with bonobos phylogenetically closest to humans, this study will examine the great apes most distanced to humans, i.e., orangutans. Orangutan personality has been primarily examined using ratings (Adams, King, & Weiss, 2012; Uher & Asendorpf, 2008; Weiss et al., 2006) and experimental

methods (Forss et al., 2015; Herrmann et al., 2011; Uher et al., 2008). The personality of orangutans has been recurrently described in five personality dimensions using the ratings from familiar keepers: extraversion, dominance, neuroticism, agreeableness and intellect (Weiss et al., 2012, 2006; Weiss & King, 2015). Additionally, Uher and colleagues (2008; 2008) combined observations and ratings and described the personality of the orangutans into 17 personality traits (e.g., aggressiveness, anxiousness, curiosity, friendliness to humans). Other studies focused only on exploratiVeness in relation with cognitive abilities (Damerius, Forss, et al., 2017; Damerius, Graber, et al., 2017; Forss et al., 2016) or social learning to determine whether more explorative rehabilitant individuals have greater chance to survive after release (Schuppli et al., 2017). To our knowledge, none of the above orangutan studies examined personality consistency using a naturalistic approach.

The current study assessed the personality consistency of rehabilitant orangutans in natural settings. More specifically, this work included juvenile, adolescent and adult Bornean orangutans at Sepilok Orangutan Rehabilitation Centre (SORC), Sabah, Malaysia. The aim of this study was to test whether naturally occurring orangutan behaviours, linked to sociability, boldness, exploratiVeness and anxiety, may be consistent across daily contexts and two time periods. We focused on these four behaviour-based personality traits as they were reliably rated (Clay et al., 2015; Freeman et al., 2013; King et al., 2008; Pederson et al., 2005; Weiss & King, 2015) and behaviourally measured by human observers (Anestis, 2005; Baker & Aureli, 1997; Koski, 2011b; Massen et al., 2013; Uher, Asendorpf, & Call, 2008). We used a coding scheme that was previously found to be reliable for measuring these four personality traits, again in a naturalistic setting but in sanctuary chimpanzees (Chapter 3); this approach was further based on multiple behavioural primate studies (e.g., Koski, 2011; Seyfarth, Silk, & Cheney, 2012; Šlipogor et

al., 2016; Uher et al., 2008). Building on our sanctuary chimpanzee study (Chapter 3) and the previous personality studies that found contextual and temporal consistency in nonhuman primates (Massen et al., 2013; Uher et al., 2008), we hypothesise that behaviours linked to personality traits in orangutans may show consistency across multiple distinct daily contexts and over time.

In addition, we compared three groups of rehabilitant orangutans in this study that differed in age and socioecological environment. On a preliminary level, it was examined to what extent these three groups differ in their pattern of contextual consistency – i.e., how consistent the orangutan behaviours are across contexts. Throughout life, orangutans seem to face different challenges. They experience developmental changes (Kuze, Malim, & Kohshima, 2005), learn new skills to survive in adulthood (Russon, 2006; van Adrichem et al., 2006) and live in different social or ecological environment (van Noordwijk & van Schaik, 2005; van Schaik, 2013). As orangutans are known to show a certain flexibility in their behaviours (van Schaik, 2013; van Schaik et al., 2016), such changes are likely to influence the behavioural expression of individuals, leading them to adjust their behaviours to some situations experienced throughout life and perhaps increase their fitness (Dingemanse & Wolf, 2013; Wolf & Weissing, 2012). In past research on nonhuman primates, age has been reported to have an effect on a personality trait score, where, for instance, young individuals tended to show higher level of boldness (Massen et al., 2013), openness/curiosity (King et al., 2008; McGuire et al., 1994) and extraversion/friendliness (Dutton, 2008; Kuhar et al., 2006; Weiss & King, 2015) than older individuals. Moreover, few empirical studies showed the influence of environmental factors (socioecological) on the expression of personality in common marmosets – *Callithrix jacchus* (Koski & Burkart, 2015; Šlipogor et al., 2016), olive baboons – *Papio anubis* (Sapolsky & Share,

2004) and chimpanzees (Bard & Gardner, 1996; Koski, 2011b; Massen & Koski, 2014). Consequently, if developmental and socioecological factors have such an influence on the behaviour of individuals, then different patterns of contextual consistency may be observed across individuals.

5.2 Method

Subjects and study site

The study includes 20 Bornean orangutan subjects at Sepilok Orangutan Rehabilitation Centre, Sabah, Malaysia: one unflanged adult male (12 years old), four adult females (mean ± SD: 21.40 ± 5.40), three adolescent males (9.94 ± 1.32), three adolescent females (9.00 ± 1.25), five juvenile males (5.40 ± 1.00), and four juvenile females (6.13 ± 0.93); see Table 5.1.

Table 5.1. Subject representation. The table shows the number of subjects and their age ranges per age category, release stage and sex group.

Age category	Release stage	Sex	Number of subjects	Age range (years)
Juvenile	Semi-released	Female	4	5-7
		Male	5	4-7
Adolescent	Released	Female	3	8-10
		Male	3	9-11
Adult	Released	Female	4	15-28
		Male	1	12

The centre is located in 43km^2 of virgin dipterocarp rainforest in the Kabili Sepilok Forest Reserve. The juvenile orangutans were semi-released. They lived in a group of peers where they had unlimited access to the forest during the day and had the possibility to stay indoors during the night. The adolescent and adult orangutans were previously released – between 1 and 20 years ago – into the forest surrounding SORC, and had,

therefore, unlimited access to the forest during day and night where they interacted occasionally with wild conspecifics. Although the orangutans partly foraged for food in the forest, they also had the chance to get supplementary food (a mix of local fruits and vegetables) at the centre in the morning (at 9:30 and 10:30 am) and afternoon (at 2:30 and 3:30 pm). None of the study subjects were related to each other. Most of the SORC orangutans were brought to the centre as orphans at the age of 1-2 years.

Data collection

Video-recordings were collected using focal animal sampling (Altmann, 1974); the subjects were followed for 3 min, once or twice a day during the morning (from 9 to 11 am) and/or afternoon (from 2 to 4:30 pm) sessions. Recording occurred over two time periods, from April to August 2016 and from end-January to July 2017. Prior to each recording session, the order of focal animal sampling was randomised to avoid any bias towards the same individual and providing a balance between time periods and contexts; hence, all orangutans were observed at different time throughout the day and in different contexts (social and non-social). A total of 42.4 hours of recording was collected for this study, with approximately 2 hours per individual (Mean ± SD = 2.12 ± 0.54 hours).

Coding scheme

The coding scheme was adapted from the study with the sanctuary chimpanzees (Chapter 3). The video-recordings obtained for this study allowed us to develop a very detailed coding scheme focusing on two parts: behavioural action and context.

From frame by frame coding of videotapes, we coded 17 behavioural actions of the focal subjects (see Table 5.2). The behavioural codes were based on previous personality

and behavioural studies in nonhuman primates which reflected to some extent four common personality traits: sociability (Ebenau et al., 2019; Eckardt et al., 2015; Freeman et al., 2013; Koski, 2011b; Neumann et al., 2013; Silk et al., 2013; Suomi et al., 1996; van Hooff, 1973), boldness (Clay et al., 2015; Dammhahn & Almeling, 2012; Freeman et al., 2013; Koski & Burkart, 2015; Nishida et al., 1999; Santillán-Doherty et al., 2010), exploratiVeness (Damerius, Graber, et al., 2017; Forss et al., 2015; Massen et al., 2013; Santillán-Doherty et al., 2010; Schuppli et al., 2017; Uher et al., 2008) and anxiety (Aureli & Waal, 1997; Baker & Aureli, 1997; Clay et al., 2015; Kalin & Shelton, 2003; Kutsukake, 2003; Kutsukake et al., 2012; Uher et al., 2008). By using video-recordings, we were able to code the behavioural actions independently from the contexts. The same action was not considered as part of the same series if there was a break of at least 5 seconds. This approach allowed to ensure the independence of the occurrences and avoid a possible inflated estimation of inter-individual differences. For each type of behavioural action, we computed frequency per hour of context. Each subject obtained one single score per behaviour for each context, and each score was standardized as a z-score.

Table 5.2. Behaviours and their definitions associated with the four personality traits of interest. The table also depicts an overview of previous studies on nonhuman primates that mentioned the different behaviours.

Personality trait	Behaviour	Definition	Studies
Sociability	Playful contact	The subject or conspecific initiates a social interaction with a conspecific or subject	(Ebenau et al., 2019; Eckardt et al., 2015; Freeman et al., 2013; Koski, 2011b; Neumann et al., 2013; Silk et al., 2013; Suomi et al., 1996; van Hooff, 1973)
	Food sharing	The subject gives or receives food from a conspecific without any resistance food (from mouth to mouth or hand to hand)	
	Food begging	The subject or conspecific is requesting food using hands or approaching his/her face towards conspecific/subject	
	Body contact	The subject or conspecific touches gently the conspecific or subject's body, presents his/her arm when approaching, embraces a conspecific or subject, presents hand or finger in conspecific or subject's mouth, the subject or conspecific's mouth enters in contact with other's mouth or with other's body part	
Boldness	Rough action	The subject hits a conspecific with hands/feet, bites a conspecific, pulls a conspecific towards him/her or pushes a conspecific away with either hands/feet, shakes an object in the direction of a conspecific	(Clay et al., 2015; Dammhahn & Almeling, 2012; Freeman et al., 2013; Santillán-Doherty et al., 2010)
	Chase	The subject follows a conspecific (walks in a direct manner or runs) while the conspecific moves away	
	Risky action	The subject grabs edible food within arms' reach from a bigger/same size conspecific who can see the subject or steals edible food from a bigger/same size conspecific who resists in some ways	
	Resist	The subject withstands the action from a conspecific	

Explorativeness	Gaze	The subject clearly directs gaze towards an object (e.g., stick, leaf, rock, rope, inedible food) or conspecific(s) (who are involved in a social interaction, eating or manipulating an object) for at least 2 seconds. Part of the subject's face needs to be visible to be counted	(Damerius, Graber, et al., 2017; Forss et al., 2015; Massen et al., 2013; Santillán-Doherty et al., 2010; Schuppli et al., 2017; Uher et al., 2008)
	Gaze Approach	The subject moves towards an object or conspecific(s) while focusing on the item. Part of the subject's face needs to be visible to be counted	
	Gaze Touch	The subject's hand or foot enters contact with an object or conspecific(s) while focusing on item. Part of the subject's face needs to be visible to be counted	
	Gaze Manipulation	The subject handles an object or join/hold conspecifics (involved in social interactions) while focusing on it	
Anxiety	Yawning	The subject opens widely his/her mouth	(Aureli & Waal, 1997; Baker & Aureli, 1997; Clay et al., 2015; Kalin & Shelton, 2003; Kutsukake, 2003; Kutsukake et al., 2012; Uher et al., 2008)
	Self-scratch	The subject rakes his/her hair with fingernails including mainly fingers or hand' s movements (gentle scratch) or including large arm movements (rough-scratch)	
	Self-touch	The subject's hand comes into contact with his/her face or body	
	Escape	The subject leaves hurriedly	
	Vigilant	The subject approaches slowly conspecific, stops occasionally to look around, grabs some food hurriedly, keeps looking around and is jumpy or freezes with no movements or vocalizations and shows tense posture	

To code the context, we used the same approach as used in Chapter 3. The context was coded frame by frame and defined according to the presence and activity of the conspecifics (not of the subject) present within 10 metres of the focal. To be considered, a context had to be displayed by at least half of the conspecifics surrounding the focal subject and last at least 10 seconds (from the first behavioural indicator defining the context displayed by the conspecifics). The duration requirement was relaxed for three of the contexts (play, aggression, and vigilance; see Table 5.3), as these three contexts could occur briefly. Because the videos were coded frame by frame, we were thus able to determine frequency and duration. With this approach, we were able to obtain a wide range of distinct contexts (see Table 5.3) which allowed us then to test for contextual consistency of the personality-linked behaviours of the orangutans.

We had a total of seven naturally occurring and distinctive contexts (see Table 5.3): aggression, feeding, vigilance, play, resting, solitude, and locomotion. Some contexts rarely occurred in our data, such as object manipulation or copulation, and were not assessed further. Furthermore, only four of the contexts (feeding, play, resting, and solitude) were included in the statistical analyses that assessed the consistency of the behaviours, as they represented at least 5% of the total duration of the complete dataset (see Table D.1). Because some of the subjects (n = 9) had lower observation time than others across all four contexts, we conducted additional statistical analyses where we considered only the three most predominant contexts (feeding, play, and solitude); this additional analysis only included 11 individuals.

Table 5.3. Definition of the analysed contexts.

Context	Definition
Aggression	The conspecifics are showing aggressive behaviours (including quarrels, fights, displays)
Feeding	The conspecifics are eating the food provided by the centre
Vigilance	The conspecifics are alert while paying attention to events happening in their surrounding
Play	The conspecifics are involved in a playful interaction
Resting	The conspecifics are sitting or lying down
Solitude	No conspecifics are around the subject
Locomotion	The conspecifics are moving on the ground, on the platform or ropes that are present in the observation area, or the trees

Two coders who were blind to the aim and hypotheses of the current study carried out the coding of behaviours and contexts. The behavioural actions were coded independently from the contexts. For inter-coder reliability purposes, 15% of the overall dataset was coded by two additional coders. Inter-coder reliability was good for both the seven contexts (Kappa = 0.75) and 17 behavioural actions (mean ICC(3,1) = 0.53, SD = 0.21, $p<0.05$).

Statistical analysis

To establish the personality of the orangutans of this study, we first examined the contextual and temporal consistency of the behaviours of the focal subject measured in the different contexts and over the two time periods. Then, all behaviours that showed consistency were subjected to a principal component analysis which allowed us to determine the personality traits that compose the orangutans' personality.

Consistency

We used the Generalizability Theory (GT) approach to assess whether the scores of the behavioural actions generalize across both contexts and time (for more information on GT, see Chapter 2; Brennan, 2011; Shavelson & Webb, 1991; Shavelson, Webb, & Rowley, 1989). Although this theory is not frequently used in animal personality studies, it previously demonstrated efficiency and robustness to assess personality in a small sample size of stumptail macaque and zebra finch (Figueredo et al., 1995), but also welfare as well as well-being ratings in chimpanzees (Robinson et al., 2017). GT allows to identify multiple sources of systematic and unsystematic errors and estimate the variance components associated with each source of variation.

We used a two-facet crossed design: subject by context by time, where subject represents the object of measurement, and context (three or four levels), as well as time (two levels), represent the facets of generalization. Additionally, we used a one-facet crossed design – subject by time – when a behaviour was observed only in one context across two time periods. For our study, we conducted a Generalizability-study (G-study) on the different facets and their interactions; a coefficient of reliability was then obtained. Two G-studies were carried out considering either the three predominant contexts (feeding, play and solitude) or all four contexts (feeding, play, resting, and solitude). Instead of having a strict cut-off to determine whether the coefficient was acceptable, we treated the coefficient per behaviour as a continuous variable to make use of most of the behaviours we measured in this study. Therefore, the behaviours showed either low, average or high consistency across context and time. We used a median split to determine which behaviours to retain for further analyses.

Determining the personality structure

Based on the G-study results when considering either three or four contexts, all behaviours that obtained a coefficient above the median (Median $_{\text{Three contexts}}$ = 0.33; M $_{\text{Four contexts}}$ = 0.27) were averaged over time and across contexts. They were then subjected to a principal component analysis (PCA); we ran two separate PCA for the two GT analyses (three or four contexts). We used eigenvalue >1 and scree-plot to determine the number of components to extract and used an orthogonal rotation (Varimax). The behaviour loadings exceeding |0.5| were considered as salient (Budaev, 2010). In order to determine which type of rotation was the most appropriate, the analysis was repeated with an oblique rotation (direct Oblimin) which allowed the components to correlate. Whether we considered either three (feeding, play, and solitude) or four contexts (feeding, play, resting, and solitude), the correlation between the two components that were extracted from the PCA was relatively low (three-context correlation: 0.17; four-context correlation: -0.06). Additionally, both rotation methods provided identical solutions regarding the behaviour loadings. Therefore, we retained and interpreted the Varimax-rotated components. Subsequently, all behaviours that loaded onto the same component (i.e., personality trait) were summed to create a single personality score for each subject.

Group comparison

We used the Kruskal-Wallis H test to compare the personality trait score of the three orangutan groups in order to assess the influence of age/socioecological factors. We calculated post-hoc analyses using Mann-Whitney U tests combined with Holm-Bonferroni correction for repeated tests. The level of significance was set at 0.05 and the tests were two-tailed (unless otherwise mentioned).

On a preliminary level, to examine to what extent age/socioecological factors could influence the contextual consistency of the orangutan personality-linked behaviours, we also descriptively compared the mean and SD of the three orangutan groups.

All analyses were computed using SPSS Statistics 25 (IBM, Chicago, IL). GT analyses were run using the syntax developed by Mushquash and O'Connor (2006). In the syntax, all negative variances were set to zero by default. However, to obtain a more accurate idea of the reliability of the dataset, we ran additional GT analyses where all negative variances were kept in the results (see Table D.2 and D.3).

5.3 Results

Assessing temporal and contextual consistency

On average, the subjects spent different amounts of time in each of the contexts for each time period (see Table D.1). Vigilance represented the lowest duration across all individuals in 2017 (0.09% of the total duration of the dataset) and did not occur in 2016 whereas feeding represented the highest duration in 2016 (46.86% of the total duration of the dataset) and solitude represented the highest duration in 2017 (40.66% of the total duration of the dataset). Across all contexts, the subjects were observed, on average, 62.99 ± 5.33 (mean ± SEM) minutes in 2016 and 64.29 ± 4.04 minutes in 2017. Over the two time periods, the subjects were observed 52.81 ± 4.13 minutes in feeding, 47.80 ± 3.42 minutes in solitude, 12.33 ± 2.10 minutes in play, and 8.42 ± 1.05 minutes in resting contexts. Additionally, the expression of the behaviours displayed within the four main contexts (feeding, solitude, play, resting) per time periods differed across the individuals (see Figure 5.1).

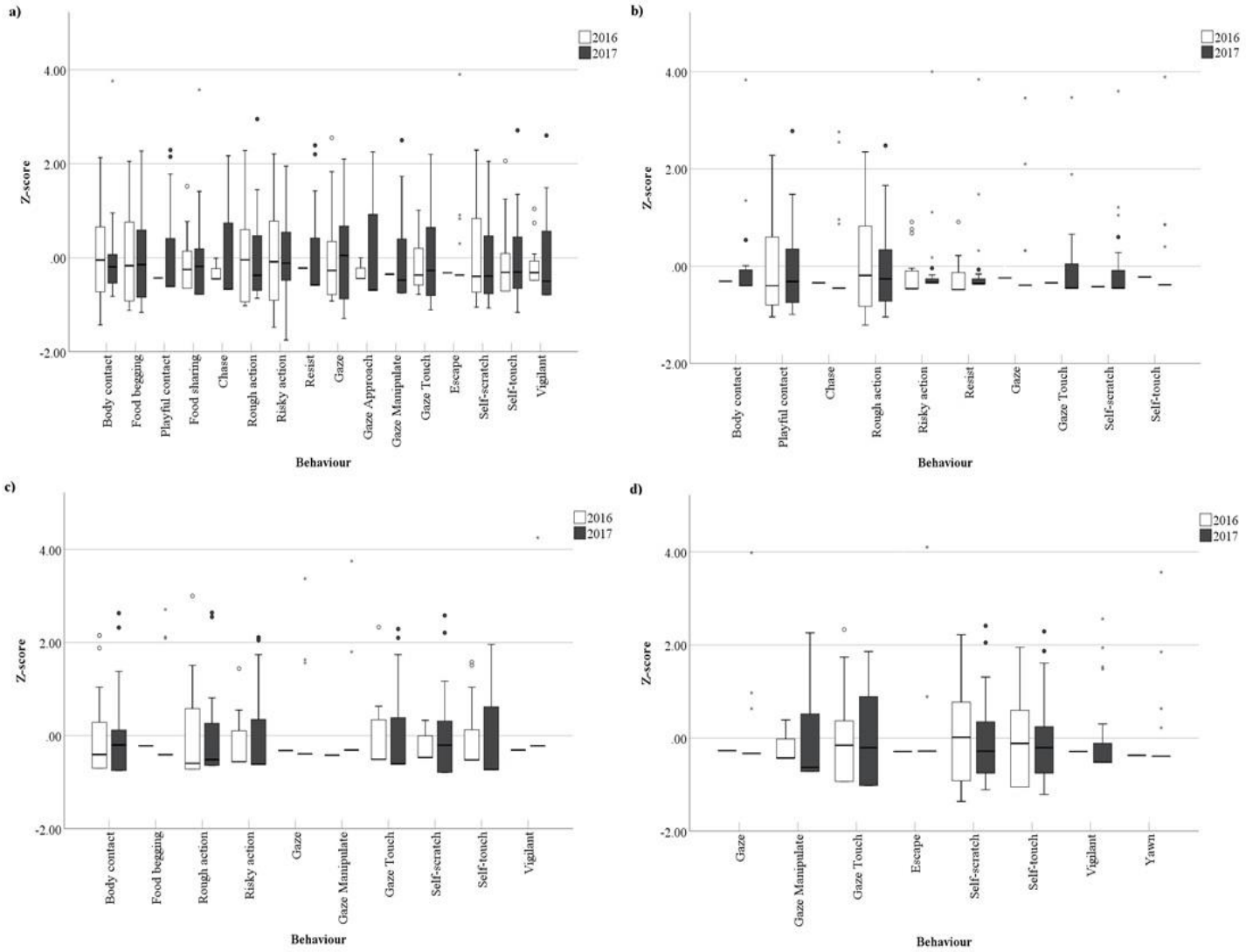


Figure 5.1. Visual representation of the behavioural expression across subjects per time period during a) Feeding, b) Play, c) Resting, and d) Solitude contexts. Data are z-scored within each time period (2016 and 2017). The thick horizontal lines indicate medians; the vertical length of the boxes corresponds to interquartile range; the thin short horizontal lines indicate the minimum and maximum values.

G study results when considering the feeding, play, and solitude contexts

G-study variance components for each of 17 behavioural action for 11 subjects are presented in Table 5.4. The other nine individuals were removed from the sample as they

had less than 10 minutes of play context. By doing so, the remaining 11 individuals were approximately observed evenly within each three contexts (see Table D.1).

Of the 17 analysed behavioural actions, four (Food begging, Food sharing, Gaze Approach, Yawn) were displayed only in one context across the two time periods of this study; Yawn did not show temporal consistency (G = 0; see Table 5.5). For Food begging behaviour, most of the variation was due to systematic differences across individuals (60.30%), whereas an important amount of residual variance (two-way subject by time interaction) was found for Food sharing behaviour (53.30%) and Gaze Approach (96.8%) indicating that individuals showed different relative score across time periods as well as other sources of error not systematically included into the G study.

Of the 13 analysed behavioural actions, the variance components attributable to subjects were 0% for three of them (Vigilant, Chase, Resist) suggesting a lack of inter-individual differences in these specific behaviours. The variance components attributable to subjects for the other 10 behavioural actions ranged from 0.60% (Gaze Manipulate) to 28.70% (Self-touch) suggesting systematic inter-individual differences. However, the variance for Gaze Manipulate was relatively low implying low variations across the individuals. Among these 10 behavioural actions, the variance attributable to the main effect for measurement contexts accounted for less than 1% of the variance for four of them (Body contact, Playful contact, Rough action, Risky action) meaning that the score for each behavioural action within each subject was consistent across contexts when averaged over subjects and time. For the other six behavioural actions, the main effect was also relatively small (less than 7%) which also indicates a fair consistency across contexts. Regarding the main effect for measurement time, the variance accounted for less than 5%

for all 10 behavioural actions indicating that the score within each subject was consistent across the two time periods when averaged over subjects and contexts. The variance attributable to the main effect for the interaction context by time accounted for less than 8% of the variance suggesting that the score for each behavioural action within each subject was consistent across both contexts and time when averaged over subjects. The variance for the subject by context and subject by time interaction ranged from 0 – 37% and 0 – 30%, respectively, indicating that there was some variability between the individuals in their patterns of contextual and temporal consistency. Finally, the three-way subject by context by time interaction (residual variance) accounted for a large percentage of variance which reflects different behavioural score per subject across time and contexts as well as other sources of error not systematically included into the G study.

Table 5.4. Estimated variance components for the G-study subject by context by time design and subject by time design per behavioural action when considering the feeding (F), play (P), solitude (S) contexts. The analysis is based on 11 subjects.

Context	F-P		F-P		F-S		F-P-S		F-P-S		F-S		F-P	
Behavioural action	Body contact		Playful contact		Escape		Self-scratch		Self-touch		Vigilant		Chase	
Effect	Var	%	Var	%	Var	%	Var	%	Var	%	Var	%	Var	%
Subject	0.02	2.90	0.11	9.90	0.27	16.60	0.09	7.40	0.56	28.70	0.00	0.00	0.00	0.00
Context	0.00	0.00	0.00	0.00	0.11	6.90	0.02	1.30	0.05	2.50	0.00	0.00	0.05	3.80
Time	0.00	0.00	0.00	0.00	0.04	2.30	0.01	1.10	0.01	0.60	0.00	0.00	0.02	1.50
Subject*Context	0.08	11.10	0.42	37.00	0.00	0.00	0.00	0.30	0.00	0.00	0.08	8.10	0.00	0.00
Subject*Time	0.16	22.60	0.31	27.80	0.00	0.00	0.00	0.00	0.00	0.00	0.07	7.30	0.30	22.40
Context*Time	0.06	8.00	0.03	3.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Subject*Context*Time	0.40	55.40	0.25	22.30	1.19	74.30	1.11	90.00	1.34	68.20	0.82	84.60	0.97	72.30

Context	F-P		F-P		F-P		F-P-S		F-P-S		F-P	
Behavioural action	Rough action		Risky action		Resist		Gaze		Gaze Touch		Gaze Manipulate	
Effect	Var	%	Var	%	Var	%	Var	%	Var	%	Var	%
Subject	0.30	28.10	0.15	14.90	0.00	0.00	0.16	10.60	0.21	14.10	0.01	0.60
Context	0.01	0.60	0.00	0.00	0.00	0.00	0.03	2.20	0.02	1.40	0.02	1.30
Time	0.00	0.00	0.00	0.00	0.00	0.00	0.05	3.60	0.00	0.00	0.00	0.00
Subject*Context	0.00	0.00	0.00	0.00	0.20	14.30	0.00	0.00	0.00	0.00	0.00	0.00
Subject*Time	0.26	24.50	0.09	9.40	0.09	6.40	0.00	0.00	0.35	24.10	0.47	30.00
Context*Time	0.06	5.90	0.07	7.30	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Subject*Context*Time	0.43	40.90	0.68	68.40	1.12	79.30	1.25	83.70	0.88	60.30	1.07	68.10

Table 5.4 (continued)

Context	F		F		S		F	
Behavioural action	Food begging		Food sharing		Yawn		Gaze Approach	
Effect	Var	%	Var	%	Var	%	Var	%
Subject	0.50	60.30	0.10	38.90	0.00	0.00	0.03	3.20
Time	0.01	0.60	0.02	7.80	0.00	0.00	0.00	0.00
Subject*Time	0.33	39.10	0.14	53.30	1.06	100.00	0.83	96.80

Var: Variance estimate was calculated using the mean squares; High values suggest that there is a large effect of the variable (see section 2.2.3, Table 2.1). %: Percentage of the total variance for each variance estimate was calculated by dividing each variance estimate by the total variance.

As the variance components attributable to subjects were 0% for four of the 17 behavioural actions, their G-coefficients were equal to 0 (see Table 5.5), suggesting that the behaviours were not consistent across contexts and time. Regarding the other 13 behavioural actions, the G-coefficients varied from 0.02 to 0.76 suggesting low to high consistency across context and time.

Following Figueredo and colleagues (1995), we constructed the G-coefficients for context and time separately for the behavioural actions that were expressed in both context and time (see Table 5.5). For context, the G-coefficients ranged from 0.21 to 1, indicating low to high consistency across feeding, play, and solitude contexts when averaged over subjects and time. Concerning time, the G-coefficients varied from 0.02 to 1, indicating low to high consistency over time when averaged over subjects and contexts.

Table 5.5. Overview of the G-coefficients for subject by context by time design as well as for context and time separately for each behavioural action when considering the feeding (F), play (P), and solitude (S) contexts.

Context	Behavioural action	G-coefficient	G-coefficient for context	G-coefficient for time
F-P	Body contact	0.09	0.21	0.11
F-P	Playful contact	0.21	0.21	0.26
F-S	Escape	0.47	1.00	1.00
F-P-S	Self-scratch	0.33	0.96	1.00
F-P-S	Self-touch	0.72	1.00	1.00
F-S	Vigilant	0.00	-	-
F-P	Chase	0.00	-	-
F-P	Rough action	0.56	1.00	0.53
F-P	Risky action	0.41	1.00	0.61
F-P	Resist	0.00	-	-
F-P-S	Gaze	0.43	1.00	1.00
F-P-S	Gaze Touch	0.39	1.00	0.37
F-P	Gaze Manipulate	0.02	1.00	0.02
F	Food begging	0.76	-	-
F	Food sharing	0.59	-	-
S	Yawn	0.00	-	-
F	Gaze Approach	0.06	-	-

The additional GT analysis which replaced the variances of 0 with the original negative values revealed a similar pattern for the different behaviours, although the values were larger overall (see Table D.2). The variance components attributable to subjects that were 0% for Vigilant, Chase, Resist and Yawning had negative values which could suggest a problem in the dataset; i.e., low amount of data point per subject which may lead to very little variation between individuals. The variance attributable to the main effect for measurement contexts and time had a small negative value (less than 0.10) for some of the behavioural actions (Body contact, Playful contact, Rough action, Risky action, Resist,

Gaze Manipulate and Gaze touch); as recommended in the literature (see Chapter 2 for details), these values can be set to zero suggesting consistency across contexts when averaged over subjects and time, or across time when averaged over subjects and contexts. With regard to the interactions between the different facets (context and time), for some of the behavioural actions, the negative values had a bigger magnitude than for the main effect measurements, ranging from 0.10 to 0.32; this suggests that we need to consider more carefully the validity of the interactions and their meaning in regards to the model due to a low amount of data per individual. Finally, the variance for the three-way interactions is relatively high for some of the behavioural actions as a result of the presence of negative values for the main effect and interaction measurements. All behaviours that were originally selected for the PCA would also be selected if using this second set of GT analysis.

Although it is important to be cautious when considering the consistency aspect of these behavioural actions for the main effects (context and time) and their interactions, it suggests a strong tendency of consistency, and it is likely that the pattern of consistency would be more substantial if there were more data per individuals.

G study results when considering the feeding, play, resting, and solitude contexts

We decided to re-run the GT analysis using all four contexts that had a total duration across the two time periods representing at least 5% of the overall total duration of the dataset (see Table D.1). This analysis was based on 20 subjects.

G-study variance components for each behavioural action are presented in Table 5.6. Of the 17 analysed behavioural actions, three (Food sharing, Gaze Approach, Yawn)

were displayed only in one context across the two time periods of this study; none of them showed temporal consistency (G = 0; see Table 5.7). The other 14 behavioural actions were displayed across two to four contexts. Among them, three (Vigilant, Chase, Resist) had a variance of 0% suggesting a lack of inter-individual differences. The variance components attributable to subjects for the other 11 behavioural actions ranged from 4.50% (Escape) to 28.20% (Self-touch) suggesting systematic inter-individual differences. Among these 11 behavioural actions, the variance attributable to the main effect for measurement contexts, time and the interaction context by time accounted for less than 1% of the variance suggesting that the score for each behavioural action within each subject was consistent across contexts and the two time periods when averaged over subjects and time or over subjects and contexts. The variance for the subject by context and subject by time interaction ranged from 0 – 24.80% and 0 – 29.50%, respectively, suggesting that there was some variability between the individuals in their pattern of contextual and temporal consistency. Finally, the three-way subject by context by time interaction (residual variance) accounted for a large percentage of variance.

Table 5.6. Estimated variance components for the G-study subject by context by time design and subject by time design per behavioural action when considering the feeding (F), play (P), resting (R), solitude (S) contexts. The analysis is based on 20 subjects.

Context	F-P-R		F-R		F-P		F-P-R-S		F-P-R-S		F-R-S		F-S	
Behavioural action	Body contact		Food begging		Playful contact		Self-scratch		Self-touch		Vigilant		Escape	
Effect	Var	%	Var	%	Var	%	Var	%	Var	%	Var	%	Var	%
Subject	0.14	11.80	0.05	5.40	0.17	17.20	0.16	16.00	0.31	28.20	0.00	0.00	0.06	4.50
Context	0.00	0.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.30
Time	0.01	0.70	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.40	0.00	0.00	0.01	0.60
Subject*Context	0.00	0.00	0.05	4.60	0.25	24.80	0.06	6.20	0.02	1.70	0.00	0.00	0.00	0.00
Subject*Time	0.00	0.00	0.17	16.60	0.16	16.10	0.04	4.00	0.00	0.00	0.32	29.50	0.00	0.00
Context*Time	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Subject*Context*Time	1.06	87.40	0.74	73.40	0.42	41.90	0.74	73.80	0.76	69.60	0.76	70.40	1.17	94.60

Context	F-P		F-P-R		F-P-R		F-P		F-P-R-S		F-R-S		F-P-R-S	
Behavioural action	Chase		Rough action		Risky action		Resist		Gaze		Gaze Manipulate		Gaze Touch	
Effect	Var	%	Var	%	Var	%	Var	%	Var	%	Var	%	Var	%
Subject	0.00	0.00	0.23	22.00	0.16	13.70	0.00	0.00	0.07	6.90	0.08	8.00	0.18	17.10
Context	0.00	0.00	0.00	0.30	0.01	0.60	0.00	0.00	0.00	0.20	0.00	0.00	0.00	0.10
Time	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.10	0.00	0.00	0.00	0.00
Subject*Context	0.00	0.00	0.00	0.00	0.00	0.00	0.06	5.20	0.00	0.00	0.06	6.20	0.00	0.00
Subject*Time	0.22	18.80	0.12	11.00	0.04	3.10	0.04	3.80	0.00	0.00	0.17	17.40	0.10	9.40
Context*Time	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Subject*Context*Time	0.93	81.20	0.71	66.80	0.95	82.60	1.05	90.90	0.99	92.90	0.69	68.40	0.75	73.40

Table 5.6 (continued)

Context	F		S		F	
Behavioural action	Food sharing		Yawn		Gaze Approach	
Effect	Var	%	Var	%	Var	%
Subject	0.00	0.00	0.00	0.00	0.00	0.00
Time	0.00	0.00	0.00	0.00	0.00	0.00
Subject*Time	1.03	100.00	1.10	100.00	1.11	100.00

As the variance components attributable to subjects were 0% for six of the behavioural actions, their G-coefficients were equal to 0 (see Table 5.7), suggesting that these behaviours were not consistent across contexts and time. Regarding the other 11 behavioural actions, the G-coefficients varied from 0.16 to 0.76 suggesting low to high consistency over both time and context.

As in the previous analysis, we constructed the G-coefficients for context and time separately for each of the 11 behavioural actions (see Table 5.7). For context, the G-coefficients ranged from 0.41 to 1, indicating average to high consistency across feeding, play, resting and solitude contexts when averaged over subjects and time. Concerning time, the G-coefficients varied from 0.25 to 1, indicating low to high consistency overall over time when averaged over subjects and contexts.

Table 5.7. Overview of the G-coefficients per behavioural action for subject by context by time design and for subject by time design when considering the feeding (F), play (P), resting (R), solitude (S) contexts. The table depicts the G-coefficients for both context and time separately.

Context	Behavioural action	G-coefficient	G-coefficient for context	G-coefficient for time
F-P-R	Body contact	0.45	1.00	1.00
F-R	Food begging	0.16	0.54	0.25
F-P	Playful contact	0.36	0.41	0.52
F-P-R-S	Self-scratch	0.56	0.72	0.80
F-P-R-S	Self-touch	0.76	0.94	1.00
F-R-S	Vigilant	0.00	-	-
F-S	Escape	0.16	1.00	1.00
F-P-R	Rough action	0.57	1.00	0.67
F-P	Chase	0.00	-	-
F-P-R	Risky action	0.47	1.00	0.81
F-P	Resist	0.00	-	-
F-P-R-S	Gaze	0.37	1.00	1.00
F-R-S	Gaze Manipulate	0.27	0.56	0.31
F-P-R-S	Gaze Touch	0.55	1.00	0.65
F	Food sharing	0.00	-	-
S	Yawning	0.00	-	-
F	Gaze Approach	0.00	-	-

As previously found for the analysis involving only three contexts, the additional GT analysis revealed a similar pattern for the different behaviours, although the values were larger overall (see Table D.3). The variance components attributable to subjects that were 0% for Vigilant, Chase, Resist, Food sharing, Yawning and Gaze Approach had negative values which could suggest a problem in the dataset; i.e., low amount of data point per subject which may lead to very little variation between individuals. The variance attributable to the main effect for measurement contexts and time had a small negative value (less than 0.10) for some of the behavioural actions (Food begging, Playful contact,

Self-scratch, Self-touch, Rough action, Risky action, Gaze Manipulate and Gaze touch). With regard to the interactions between the different facets (context and time), for some of the behavioural actions (Body contact, Escape, Risky action), the negative values had a bigger magnitude than for the main effect measurements, ranging from 0.14 to 0.18. Finally, the variance for the three-way interactions is relatively high for some of the behavioural actions as a result of the presence of negative values for the main effect and interaction measurements. All behaviours that were originally selected for the PCA would also be selected if using this second set of GT analysis.

As mentioned in the previous analysis, the interpretation of the consistency for some of the behaviours has to be considered carefully; however, a bigger dataset would provide a stronger pattern of consistency for most of the behavioural actions.

Personality structure

Based on GT including feeding, play and solitude

We retained all behavioural actions that had a G-coefficient above the median (M = 0.33); hence, nine behavioural actions were retained for the PCA. By including these nine behavioural actions, both Kaiser–Meyer–Olkin (KMO = 0.412) and Bartlett's Test of Sphericity ($p > 0.001$) were not acceptable, indicating poor suitability of the data (Budaev, 2010). Consequently, we removed two additional behavioural actions (Escape and Self-touch) from the sample. Escape was displayed in both feeding and solitude contexts; however, fleeing in solitude context might be difficult to interpret and misleading as we cannot be certain for the reasons for such behaviour in this context. Self-touch was

removed based on its individual KMO which was relatively low suggesting its unsuitability for the analysis. Therefore, we retained seven behavioural actions for the final PCA. The Kaiser–Meyer–Olkin (KMO = 0.663) and Bartlett's Test of Sphericity ($p < 0.05$) indicated the suitability of the data. We extracted two components which explained 72% of the variance (see Table 8). The first component explained 52% of the variance and had positive loadings of six behavioural actions that related to social, bold and explorative actions; this component was therefore labelled "exploratiVeness-boldness-sociability". The second component accounted for 20% and had positive loadings of one behaviour that related to anxious behavioural actions; we decided to label it "anxiety".

Table 5.8. Behavioural action loadings after Varimax rotation when considering the feeding, play and solitude contexts.

Behavioural action	ExploratiVeness-Boldness-Sociability	Anxiety
Self-scratch		0.910
Gaze	0.903	
Gaze Touch	0.842	
Risky action	0.744	
Rough action	0.669	-0.623
Food begging	0.731	
Food sharing	0.745	

Only loadings above |0.5| are reported.

Based on GT including all four contexts (feeding, play, resting and solitude)

We decided to keep all behavioural actions that had a G-coefficient above the median (M = 0.27); hence, nine behavioural actions were retained for the PCA. The Kaiser–Meyer–Olkin (KMO = 0.749) and Bartlett's Test of Sphericity ($p < 0.001$) indicated the suitability of the data. We extracted two components which explained 71% of the variance (see Table

9). The first component explained 51% of the variance and had positive loadings of six behaviours related to social, bold and explorative actions; this component was therefore labelled "sociability-boldness-explorativeness". The second component accounted for 20% and had positive loadings of four behaviours that related to social, anxious and bold actions. Therefore, the second component was labelled "sociability-anxiety".

Table 5.9. Behavioural action loadings after Varimax rotation when considering the feeding, play, resting, and solitude contexts.

Behavioural action	Sociability-Boldness-Explorativeness	Sociability-Anxiety
Body contact		0.690
Playful contact	0.915	
Self-scratch		0.810
Self-touch		0.563
Rough action	0.926	
Gaze	0.895	
Gaze Manipulate	0.863	
Gaze Touch	0.716	
Risky action	0.538*	0.598

*Because Risky action loaded higher on component 1 than component 2 with an Oblimin rotation, we decided to retain its loading on the first component to facilitate the interpretation. Only loadings above |0.5| are reported.

Influence of age/socioecological factors on personality trait scores

Only the sociability-boldness-exploratiVeness trait (obtained with the PCA considering the four contexts) was significantly different between the three orangutan groups of this study, $H(2) = 9.6$, $p < 0.01$ (see Table 5.10). The post-hoc analyses using Mann-Whitney U test with Holm-Bonferroni correction revealed that the juvenile orangutans showed a

significantly higher level of sociability-boldness-explorativeness trait (Mean ± SD = 0.55 ± 0.72) than adult orangutans (-0.97 ± 0.30), $z = -2.87$, $N_{Juveniles} = 9$, $N_{Adults} = 5$, $p < 0.01$.

Table 5.10. Age/socioecological group analysis. Overview of the statistical analyses assessing the age/socioecological group effect on the personality trait score using Kruskal-Wallis H. The age group was determined based on the median age across the two time periods. The analysis included 20 subjects. The post-hoc analysis using the Mann-Whitney U test is also presented for each age category comparison.

			Juvenile vs Adolescent			Juvenile vs Adult			Adolescent vs Adult		
Personality trait	Kruskal-Wallis H	Exact Sig.	Mann-Whitney U	Z	Exact sig. (2-tailed)	Mann-Whitney U	Z	Exact sig. (2-tailed)	Mann-Whitney U	Z	Exact sig. (2-tailed)
Sociability-boldness-exploratiVeness	9.60	< 0.01	15.00	-1.41	0.181	1.00	-2.87	< 0.01	4.00	-2.01	0.052
Sociability-Anxiety	1.72	0.442	-	-	-	-	-	-	-	-	-

Influence of age/socioecological factors on contextual consistency

We ran a GT analysis for each behavioural action per age category (juvenile, adolescent, adult) in order to see whether the behaviour consistency differs across the age/social groups (see Table D.4). Table 5.11 shows the overall G coefficient as well as the G coefficient calculated for the context only.

If we look at the behavioural actions per personality category, juveniles showed higher level of consistency for social behavioural actions (Mean ± SD = 0.72 ± 0.48) compared to adolescents (0.16 ± 0.29) and adults (0.33 ± 0.58). A similar pattern was observed for anxiety-related behavioural actions (Juvenile: 0.72 ± 0.48; Adolescent: 0.29 ± 0.32; Adult: 0.25 ± 0.50). For boldness-related behavioural actions, the pattern changed; both adolescents (0.75± 0.50) and adults (1.00 ± 0.00) showed higher level of contextual consistency than juveniles (0.33 ± 0.58). Finally, both juveniles (1.00 ± 0.00) and adults (0.92 ± 0.12) showed higher level of contextual consistency compared to adolescents (0.35 ± 0.50) regarding the exploratiVeness-related behavioural actions.

Table 5.11. Overview of the G-coefficients for subject by context by time design as well as for context per age category. We considered the feeding (F), play (P), resting (R), and solitude (S) contexts.

			G-coefficient overall			G-coefficient for context		
Personality trait	Context	Behavioural action	Juvenile	Adolescent	Adult	Juvenile	Adolescent	Adult
Sociability	F-P-R	Body contact	0.40	0.27	0.00	1.00	0.49	0.00
	F-R	Food begging	0.23	0.00	0.68	1.00	0.00	1.00
	F-P	Playful contact	0.12	0.00	0.00	0.17	0.00	0.00
Anxiety	F-S	Escape	0.45	0.00	0.00	1.00	0.00	0.00
	F-P-R-S	Self-scratch	0.68	0.47	0.00	1.00	0.63	0.00
	F-P-R-S	Self-touch	0.75	0.31	0.89	0.88	0.25	1.00
	F-R-S	Vigilant	0.00	0.00	0.00	0.00	NA	0.00
Boldness	F-P	Chase	0.00	0.00	0.44	0.00	0.00	1.00
	F-P-R	Rough action	0.00	0.79	0.15	NA	1.00	1.00
	F-P	Resist	0.00	0.67	0.09	0.00	1.00	1.00
	F-P-R	Risky action	0.01	0.11	0.00	1.00	1.00	NA
Exploratiiveness	F-P-R-S	Gaze	0.39	0.00	0.19	1.00	NA	0.83
	F-R-S	Gaze Manipulate	0.36	0.00	0.00	1.00	0.00	NA
	F-P-R-S	Gaze Touch	0.30	0.54	0.27	1.00	0.70	1.00

5.4 Discussion

The current study tested for the contextual and temporal consistency of personality-linked behaviours of rehabilitant orangutans using naturalistic observations. Eleven behavioural actions were found to be consistent, with nine of them showing a high level of consistency across two to four distinct contexts (feeding, play, resting, and solitude) and over time (5 months apart). These nine behaviours loaded onto two components: a combination of sociability/boldness/exploratiiveness, and anxiety. A similar personality structure (i.e., two traits constituted of similar behavioural indicators) was observed when considering the three most predominant contexts (feeding, play, and solitude), suggesting that the personality profile of these orangutans is relatively stable. The current study, thus, provides

the first evidence of contextual and temporal consistency in orangutans' behaviours using a naturalistic approach. To our knowledge, temporal and contextual consistency of orangutan personality has only been demonstrated in zoo subjects tested in an experimental setting (Uher et al., 2008). Besides orangutans, temporal (Koski, 2011; Chapter 3) and contextual (Chapter 3) consistency of personality, measured using naturalistic observations, have been only found in one another great ape species, chimpanzees. Hence, we could speculate that this characteristic of consistency may be also observable in bonobos and gorillas when measured in their day-to-day life since consistency was found for the most distantly and closest related great ape species. Using a naturalistic approach which measures personality-linked behaviours in a natural setting can improve our understanding of how nonhuman primates express their personality in their daily lives.

The orangutans we studied showed consistencies across the examined contexts, but this consistency depended on the behavioural actions and the contexts in which they were measured. For instance, some of the exploratiVeness-related behaviours (e.g., Gaze Touch) showed high consistency across feeding, play, resting and solitude whereas others (e.g., Gaze) showed low consistency across the same contexts. However, this pattern changed when examining the consistency of these behaviours across feeding, play and solitude only. Similarly, for sanctuary chimpanzees, we found that some personality-linked behaviours differed in their patterns of consistency, where some behaviours, but not all, showed consistency across all contexts (Chapter 3). These different patterns among the behavioural actions may be explained in that the settings affect to some extent the expression of the personality-linked behaviours. Such effects may result in higher level of consistency between some contexts as previously observed in an experimental study in common marmosets (Koski & Burkart, 2015). It is important for future research to further

consider multiple distinct contexts when examining personality. Each context is likely to convey different relevant information about the personality structure of individuals who may express their behaviours according to the context experienced. Then, focusing on various contexts may help reveal subtle inter-individual differences (Chapter 3) and cover different facets of expression of a given personality trait (Carter et al., 2012a).

With regards to the temporal consistency, regardless of the contexts, most the behavioural actions showed high consistency over short periods of time (5-month break). The finding supports a previous study on zoo orangutans that also found temporal stability in an experimental setting (Uher et al., 2008) and is comparable to the findings on other nonhuman primates where temporal consistencies were found when behaviours related to sociability, boldness, exploratiVeness and anxiety were measured in natural settings (Brent et al., 2014; Byrne & Suomi, 1995; Koski, 2011b; Maestripieri, 2000; Seyfarth et al., 2012; Suomi et al., 1996; von Borell et al., 2016). Temporal consistency of these personality-linked behaviours certainly allows us to consider that they are enduring characteristics of the individuals' personality. To fully understand how nonhuman primates express their personality-linked behaviours on a day-to-day basis, it is important to measure their behaviours across multiple periods of time, thus we can provide a detailed description of their personality profile

Moreover, individual orangutans may show different levels of temporal and/or contextual consistency. Depending on the behavioural action measured in this study, the GT analyses revealed that the individuals differed in their level of contextual and temporal consistency. These findings imply that some of the orangutans might be more consistent across contexts and/or time in their personality-linked behaviours than others, matching

what was found previously in orangutans, as well as in the other three extant great ape species (Uher et al., 2008; Chapter 2). Whereas some individuals might show consistency in their personality-linked behaviours across contexts and over time throughout their lives, others might instead be more malleable and sensitive to changes reflecting individuals' ability to easily adapt to situation or context over time (Dingemanse, Kazem, Réale, & Wright, 2010). These different strategies, whether being consistent or plastic in the expression of personality-based behaviours, are likely to have an impact on the fitness of individuals (e.g., reproductive success: Réale, Gallant, Leblanc, & Festa-Bianchet, 2000; life span: Altschul et al., 2018).

Furthermore, the current study compared the personality scores across three orangutan groups, differing on the age and socioecological level. The results of this work suggest that the juvenile orangutans may have shown a significantly higher level of the combined sociability-boldness-explorativeness trait than the adult orangutans. Additionally, our findings preliminary showed different patterns of contextual consistency in the different personality-linked behaviours across the three groups. Age differences were previously reported in great ape species for each of these separate personality traits (King et al., 2008; Massen et al., 2013; Weiss & King, 2015), however, here, we cannot ignore the influence of others factors on our data. Past research in nonhuman primates showed group differences in the expression of some personality traits, such as sociability, exploration and boldness (Bard & Gardner, 1996; Koski, 2011b; Koski & Burkart, 2015; Šlipogor et al., 2016) and personality adjustment from individuals immigrating to a new group of primates (Sapolsky & Share, 2004). These findings suggest that the socioecological environment may influence the personality expression. The juvenile orangutans living in the rehabilitation centre still depended on human care and were still in

need to learn survival skills prior to their full releases (Russon, 2006; van Adrichem et al., 2006; van Noordwijk & van Schaik, 2005). In contrast, the adolescent and adult orangutans were free-ranging in the forest surrounding SORC. Throughout life, individuals might find the need to readjust their behaviours depending on the situation they face (Dingemanse & Wolf, 2013) to maximise their benefits in terms of fitness (Wolf & Weissing, 2010, 2012). Different social group constellation or rearing experience (Bard & Gardner, 1996; Koski & Burkart, 2015; Schuppli et al., 2017), as well as different stages of rehabilitation (Damerius, Graber, et al., 2017), are likely to have a different impact on an individual‘s behavioural expression in the course of their lives. Future research should examine further the impact of these different factors on the consistency of personality traits, which would provide new insight into the understanding of personality in nonhuman primates.

It is, however, important to remain cautious in the interpretation of the results, as our study also presents some limitations. Using behavioural observations in natural setting has great advantages in terms of ecological validity (Freeman et al., 2011) but also has some drawback in terms of the amount of data obtained per subject. Here, the overall dataset used to run the analyses was relatively small (i.e., about 42h of recordings) in comparison to past personality research in nonhuman primates [e.g., over 100 hours of recording: (Anestis, 2005; Castles et al., 1999; Ebenau et al., 2019; Tkaczynski et al., 2018)]. It is possible that the amount of data per individual is not sufficient enough to provide a clear picture of their personality profile. We tried to adopt a cautious approach to analyse our data by focusing, first, on the three predominant contexts where most of the data were available, and we, then, run a second set of analysis by considering all relevant contexts. We used GT to assess behavioural consistency which is a statistical framework that allowed us to breakdown the different sources of errors coming potentially from the

contexts, time and their interactions. Nonetheless, by looking at the interactions, we, therefore, reduced even more our dataset available which resulted in negative variances for some of the behavioural actions. Such negative values are reported to indicate a sampling error (Shavelson & Webb, 1991). However, in case of small negative variances (very close to 0), it is widely recommended to set the variances to 0 (Brennan, 2001), which, to some extent, suggests a lack of variation across context and/or time (i.e., consistency). Although the findings presented in this study have to be carefully considered, it is likely that if the individuals had more data points per context and per time period, we would obtain stronger patterns of behavioural consistency.

In summary, orangutans showed both contextual and temporal consistency of their personality-based behaviours when measured in multiple distinct contexts. Furthermore, as previously shown in sanctuary chimpanzees (Chapter 3), the orangutans in our study showed hereby different patterns of consistency in their behaviours. It is important for future research on nonhuman primates to acknowledge the context when assessing a personality trait using naturalistic observations, as different patterns of consistency are likely to occur across individuals. By doing so, subtle inter-individual differences can be revealed in the data. Nevertheless, in this study, we also show that personality appears to be relatively plastic, in that the individuals may show a certain consistency throughout their life, but they may also demonstrate a certain malleability in their behavioural expression to adjust to the environment they evolve in and the experience acquired. A personality trait can be expressed by different behaviours in various day-to-day contexts throughout a lifetime, highlighting the complexity of measuring personality. To further understand this complexity, future studies could perhaps define the personality traits on a more conceptual level where diverse behaviours are selected to generate the different

personality traits of interest rather than focusing only on consistent behaviours (Uher, 2008a; Uher et al., 2008; Uher, Werner, et al., 2013). Using naturalistic observations, in combination perhaps with other methods of assessments (Chapter 4), are likely to help obtain a detailed description of personality structure and possibly develop further our understanding of the complexity of personality. Since the naturalistic approach relies on multiple behaviours that occur naturally in the daily life of individuals, such an approach could be beneficial for researchers to use to study the personality of free-ranging wildlife, like the rehabilitant orangutans of this study.

6. General Discussion

In this doctoral thesis, I examined the concept of personality in chimpanzees and orangutans, the closest and most distanced great ape species relative to humans, respectively. My particular focus was on two key aspects of personality, contextual consistency and temporal consistency, and my primary method was naturalistic observations. My results showed that multiple behaviours reflected the personality traits of sociability, boldness, explorativeness and anxiety in the sanctuary chimpanzees and the rehabilitant orangutans of my work. The contexts were ecologically relevant to the two species as they naturally occurred in the daily life of the individuals (e.g., feeding, resting or play). The chimpanzees' personality assessment was further explored by comparing different methods (i.e., ratings, experimental and naturalistic observations) with a goal to understand the contribution of each method for assessing the four personality traits.

In this final chapter, I discuss the main findings from my empirical studies and the implications in relation to the animal personality research. I present the major strengths as well as limitations of my work and suggestions for future research.

6.1 Discussion of main findings and their implications

6.1.1. *Benefits of using naturalistic observations*

Along with other methods of assessment, to understand individuals' personality, it is important to observe them in the course of their lives where they can interact naturally with their environment. By avoiding any intervention from the observer, we are able to access real-life data which are likely to reflect a true-to-life nature of personality (McDonald,

2008). Indeed, the individuals of this research expressed their personality using multiple behaviours occurring naturally in multiple situations and contexts throughout time. Naturalistic observations provided, therefore, the possibility to capture this diversity of expression of the different personality traits as it has been previously demonstrated in both humans (Mehl et al., 2006) and nonhuman primate species (Brent et al., 2014; Byrne & Suomi, 1995; Koski, 2011b; Seyfarth et al., 2012).

Here, multiple benefits have been highlighted throughout this thesis, regarding both the methodologies used in this work (see Chapter 2 for a review) and within the three empirical studies (Chapters 3, 4, and 5). First, we were able to target multiple behaviours that were linked to the four personality traits of interest: sociability, boldness, exploratiVeness and anxiety. We measured these behaviours in multiple ecologically relevant contexts that included multiple situations across multiple occasions allowing us to cover multiple facets of expression of the same trait. Second, because we observed this diversity of expression of personality trait, we were able to more clearly show subtle inter-individual differences in terms of contextual and temporal consistency. Finally, using this approach allowed us to unravel further the complexity of measuring personality consistency in the studied nonhuman primates. In particular, this complexity was based on three levels: behaviour, context and time.

The naturalistic method, especially as it was combined with video-recordings, allowed us to use a very detailed approach to measure the four personality traits of interest in the sanctuary chimpanzees and rehabilitant orangutans. The context was independently coded from the behaviours and two behavioural actions were part of the same series if they occurred within five seconds of one another. By doing so, we were able to ensure the

independence of the behaviour occurrences and avoid a possible inflated estimation of inter-individual differences. This method contributed to a detailed description of an individual's personality traits in relation to the contexts they occurred in. The personality-linked behaviours were consistent across two to five contexts (feeding, grooming, play, resting, and solitude) and over the two time periods. As previously found in behavioural and personality studies on nonhuman primates [sociability: (Eckardt et al., 2015; Koski, 2011b; Stoinski, Kuhar, Lukas, & Maple, 2004); boldness: (Dammhahn & Almeling, 2012; Massen et al., 2013; Santillán-Doherty et al., 2010); explorativeness: (Massen et al., 2013; Schuppli et al., 2017; Uher et al., 2008); anxiety: (Baker & Aureli, 1997; Kutsukake et al., 2012; Uher et al., 2008)], the selected and measured behaviours represented the four personality traits of interest as revealed with the principal component analysis.

To refer to personality, highlighting individuality in the expression of personality traits is crucial (Allport, 1961; Gosling, 2001, 2008). Here, both chimpanzee and orangutan subjects showed intra-individual consistency in their personality-based behaviours across multiple contexts and time while maintaining their inter-individual variations. These results are in line with past nonhuman primate personality studies (Ebenau et al., 2019; Koski, 2011b; Massen et al., 2013; Seyfarth et al., 2012; Šlipogor et al., 2016; Tomassetti et al., 2019; Uher et al., 2008; Weiss et al., 2015). Our findings suggest that both great ape species express their personality in their day-to-day life and the naturalistic observations used in this work allowed us to reveal their inter-individuality.

Among our data, we found that some of the individuals showed higher levels of temporal and/or contextual consistency than others in the expression of their personality traits. Some individuals may have expressed their personality in contexts that differ from

conspecifics, as revealed in the generalizability theory (GT) analyses. Different strategies of consistency may be adopted by the individuals to maximise their individual fitness (Wolf and Weissing 2012), such as an increase of the reproductive success (Réale, Gallant, Leblanc, & Festa-Bianchet, 2000) or lifespan (Altschul et al., 2018). Because the personality assessment used in this research allowed us to capture this diversity of expression of the different personality traits, we were able to reveal subtle inter-individual differences in terms of personality consistency.

Individuals can express their personality in different ways throughout life, and if we want to grasp this diversity of expression, it is important to use a method that allows to achieve this goal. Based on the findings of this research, the naturalistic approach provided us with a broader perspective on the four personality traits of interest in comparison with the other two methods of personality assessment, and particularly experimental observations which measured boldness, explorativeness and anxiety in two specific situations (threatening-like situation and novel situation). Of course, it is important to note that the rating method that we used provided similar views regarding the four traits because the raters were asked to make their judgements based on their overall idea of the chimpanzees' behaviour occurring over time and across multiple contexts. However, people tend to recall more major events rather than multiple day-to-day events (Koppel et al., 2013), so we cannot ignore the possibility that the ratings of the keepers were perhaps influenced by past major events that occurred in the focal individual's life (e.g., fight).

Overall, the naturalistic approach was evidently appropriate to measure both temporal and contextual consistency of these four personality traits in the day-to-day life of the two great ape species. Considering multiple behaviours (e.g., groom, gaze, risky action,

self-scratch) and contexts (e.g., feeding, social, resting, solitude) over multiple occasions contributed to a better understanding of the complexity of expression of these four personality traits of the chimpanzees and orangutans of this work. By using this approach, we were able to cover different facets of expression of the same personality trait and revealed interesting inter-individual differences.

6.1.2. *Context is a key criterion in personality*

Past research primarily focused on measuring personality consistency over time using ratings (Freeman et al., 2013; Uher & Asendorpf, 2008; Weiss et al., 2017), experimental observations (Kutsukake et al., 2012; Uher & Asendorpf, 2008; Uher et al., 2008) and naturalistic observations (Byrne & Suomi, 1995; Koski, 2011b; Seyfarth et al., 2012). For a long period of time, it was believed that personality was not consistent across contexts (or very little); this observation mostly stemmed from the human literature (Kenrick & Funder, 1988; Mischel, 1968). However, empirical evidence showing contextual consistency started to emerge in both human (Funder & Colvin, 1991; Furr & Funder, 2004) and nonhuman (Schuster et al., 2017; Uher et al., 2008) literature. More specifically, contextual consistency of personality in nonhuman primates has been demonstrated using only experimental observations (e.g., Dammhahn & Almeling, 2012; Koski & Burkart, 2015; Sussman & Ha, 2011; Tkaczynski et al., 2018). Here, for the first time in nonhuman primates, we found evidence for contextual consistency of personality in both sanctuary chimpanzees and rehabilitant orangutans using naturalistic observations. This finding contributes greatly to the enrichment of the field. Indeed, apart from being measurable in experimental settings (e.g., Massen et al., 2013; Uher et al., 2008), we demonstrated that

the contextual consistency aspect of personality can be established in the day-to-day life of two great ape species.

The subjects of this work showed consistency in their personality-related behaviours across two to four different contexts. Unlike previous experimental studies that showed contextual consistency only in specific situations [e.g., small vs big novel objects to measure explorativeness: (Koski & Burkart, 2015); snake vs bird situations to measure boldness: (Koski & Burkart, 2015)], here, we actually targeted multiple distinct contexts. A context can include multiple specific situations (Sih, Bell, & Johnson, 2004). For instance, a feeding context can be both with and without predators (Sih, Bell, & Johnson, 2004) or a vigilant context can be related to both food anticipation (Waitt & Buchanan-Smith, 2001) or background disturbance (Kutsukake, 2006; Shultz et al., 2003). By focusing on multiple situations and multiple contexts, we were able to provide a detailed overview of the four personality traits of interest in the everyday life of the subjects.

Furthermore, having such a range of contexts allowed us to highlight different patterns of consistency in the personality-based behaviours (Chapters 3 and 5). Two different patterns emerged in our data. On one hand, some of the behaviours showed consistency across all four contexts (e.g., self-scratch) whereas others were consistent only across two contexts (e.g., escape). These different patterns observed among the behaviours may be explained in that the contexts perhaps affected to some extent the expression of the personality-linked behaviours. Thus, it may have resulted in a higher level of consistency between some contexts compared to other contexts as previously observed in an experimental study in common marmosets (Koski & Burkart, 2015). On the other hand, some of the behaviours (e.g., gaze touch) showed a high level of consistency in all four

contexts whereas other behaviours (e.g., gaze) showed a lower level of consistency in the same contexts. In human personality research, it was suggested that the consistency of some behaviours may result from different causal mechanisms (Funder, 2006; Furr & Funder, 2004). More specifically, automatic behaviours (e.g., body posture) tend to show high consistency across situations whereas controlled behaviours (e.g., verbal behaviours) tend to depend on the situation (Weisbuch et al., 2010). Consequently, it is possible that the different patterns of consistency observed in our data may result from the specific type of behaviour (automatic vs controlled), although we cannot confirm with certainty.

Personality traits can be expressed by different behaviours in different contexts and either or both of these can vary throughout a lifetime, highlighting the complexity in measuring personality. By summarizing over the contexts (i.e., not specifying the context) when assessing personality consistency, past research might have missed some relevant information in terms of determining the personality structure of a species (e.g., Brent et al., 2014; Ebenau et al., 2019; Suomi et al., 1996). Capturing a wide range of distinctive contexts, by including the ones that show different patterns of personality-based behaviours, is highly important for understanding the personality trait of interest – targeting different facets of expression of the same personality trait. Both time and context are key criteria to test when examining personality (Gosling, 2008; Sih, Bell, Johnson, et al., 2004), so it seems essential to consider both of them in order to obtain a more detailed overview of the personality traits examined. The context seems to play a key role in defining personality, and by using this information, we can perhaps understand better how the different behaviours that constitute the different traits work together and how they are expressed in the daily life of individuals. By focusing on context, we were able to capture

the complexity of personality and provide a detailed picture of its structure in the day-to-day life of the two species studied.

6.1.3. Importance of multi-methods

The personality structure of an individual is characterised by a combination of traits that are dynamically organised with one another (Allport, 1961), and each personality trait can be composed of a variety of related but different behaviours. In this study, boldness was, for instance, measured by rough actions displayed in a play interaction and stealing food from a conspecific in a feeding context whereas anxiety was measured by self-scratching in an aggression context and vigilant behaviour in a solitude context (Chapters 3, 4 and 5). Previous research reported that some personality traits may be divided into multiple sub-traits (facets) (Carter et al., 2012a; Koski, 2011a; Tkaczynski et al., 2018). For instance, sociability can be characterised by tactility, equitability, positive affect and a general sociability trait (Koski, 2011a; Tkaczynski et al., 2018). Whereas tactility is associated with a grooming context (Tkaczynski et al., 2018), positive affect is related to a play context (Koski 2011b). Because personality traits can be expressed in different ways, they are likely to be composed of multiple facets, so it is important to use methods that can target these different ways of expression.

By systematically comparing the three methods of assessment, each of which measured the same personality trait, we found mix findings in terms of coherence across the methods. More specifically, after running successive pairwise comparisons across the methods on the trait level, we found two significant positive correlations between the two formats of ratings for boldness, and between the rating (adjective) and experimental observations for exploratiVeness. Additionally, we found a positive trend for sociability

when comparing the rating method (adjective) with naturalistic observations. These findings are in line with previous research in nonhuman primates which found similar observations for boldness (Uher, Werner, et al., 2013), explorativeness (Uher & Asendorpf, 2008) and sociability (Freeman et al., 2013; Tkaczynski et al., 2018; Uher & Asendorpf, 2008; Uher & Visalberghi, 2016). Here, the coherence observed between the methods for these specific personality traits demonstrates their interchangeable nature as they showed high similarities in their assessment. More specifically, boldness seems to be reliably assessed with both a series of adjectives and behavioural descriptors, explorativeness measured in a novel object test and sociability measured in day-to-day contexts reflect both the raters' judgements when using adjectives. However, unlike previous research (Uher & Asendorpf, 2008; Uher, Werner, et al., 2013), we found two negative trends for anxiety when comparing the rating (adjective) method with naturalistic observations, and the naturalistic observations with the experimental observations. It is possible that what the keepers perceived as anxiety in the questionnaire and what was measured in the experimental situation (snake test) reflected a different type of anxiety in the everyday lives of the individuals. These results suggest that the methods did not measure the same facet of the personality trait.

Furthermore, we compared the naturalistic observations with the rating (behavioural descriptor) on the behaviour level to provide a more detailed picture of what each of these two methods measures. Here, we found positive relationships for sociability, boldness and anxiety. For instance, for boldness-related item/behaviour, an individual who was rated high on "When playing with stronger chimps, *X* plays rough and chases them" was also observed playing rough and chasing individuals in a paly context when measured with naturalistic observations. This finding clearly suggests that we can use either one

because they are highly similar in assessing this specific facet of boldness. In addition, we found a significant negative relationship for one of the sociability-related item/behaviour. One reason that could explain this lack of coherence is the choice of the behavioural indicator to reflect the item ("When relaxing, *X* is nearby other chimps"); I used body contact as an indicator of sociability. However, it is possible that a measurement of social proximity would be more accurate to match the item.

Overall, the comparison between the methods suggests that each may capture a different set of information which allows to better grasp the diversity of expression of each of the four personality traits. Of course, depending on the trait, it is important to note that ratings and experimental observations might not always have an equivalent behaviour when measured with naturalistic observations which may explain the lack of coherence between methods which was previously reported in the literature (Neumann et al., 2013; Tkaczynski et al., 2018). The discrepancy observed in our findings testifies to the difficulties to assess personality in nonhuman primates. Certainly, using all three methods to assess personality would be a huge endeavour to carry out in any animal personality study, but using perhaps one method in combination with another could help obtain a finer description of the personality trait targeted and provide a better understanding of the complexity of personality.

6.1.4. Intrinsic and extrinsic influences

Both intrinsic (e.g., genetic, age, sex) and extrinsic (e.g., rearing experience, social group constellation) factors are known to influence the expression of a personality trait (Bard & Gardner, 1996; Brent et al., 2014; Carter et al., 2014; Martin & Suarez, 2017). Here, we preliminary investigated the potential influence of age and environment (social and

ecological) on the contextual consistency of the behaviours of rehabilitant orangutans (Chapter 5). Although it is impossible to make strong claims (as no inferential statistics were conducted) and determine the exact influence (age, social, ecological), the orangutans of this study seemed to show different patterns of contextual consistency in the different categories of behaviours measured. To date, considering only age and socioecological factors, past research in nonhuman primates only found evidence of the influence of the age (King et al., 2008; Massen et al., 2013; Weiss & King, 2015) or social environment (Bard & Gardner, 1996; Damerius, Graber, et al., 2017; Koski & Burkart, 2015; Sapolsky & Share, 2004) on the expression of personality. This contextual variability observed across rehabilitant orangutans may suggest that the individuals are able to show some plasticity in the expression of their behaviours, allowing them, in return, to adjust to the situation experienced and increase their fitness (Dingemanse & Wolf, 2013; Wolf & Weissing, 2012). These different effects (age, social, ecological) might all have played a role to influence the expression of the personality-linked behaviours of these rehabilitant orangutans. A change in the social constellation or a different rearing experience (Bard & Gardner, 1996; Koski & Burkart, 2015; Schuppli et al., 2017), as well as different stages of rehabilitation (Damerius, Graber, et al., 2017), have been reported to affect the behaviours of individuals, so it should not be surprising to observe different patterns of personality consistency throughout an individual's life.

6.2 Methodological considerations and future directions

6.2.1. Strategy to adopt with small dataset

As previously pointed out, a naturalistic approach has some limitations in terms of collecting enough data per individual (see Chapter 2; Freeman & Gosling, 2010; Freeman et al., 2011). Here, we dealt with smaller dataset (e.g., 42-53 hours of recording) in comparison with previous personality studies in primates which often use larger dataset (e.g., over 100 hours of recording) to determine the personality structure of the species studied (Anestis, 2005; Castles, Whiten, & Aureli, 1999; Ebenau et al., 2019; Konečná et al., 2008; Koski, 2011b; Tkaczynski et al., 2018). However, to counteract this issue, the analyses were split into two parts: First, the consistency was assessed across the most predominant contexts (feeding and resting for the chimpanzees; feeding, play and solitude for the orangutans) and, second, the consistency was reassessed across all relevant contexts. Regardless of the approach used to analyse the data, the GT analyses revealed similarly consistent behaviours, leading to stable solutions when determining the personality structure of chimpanzees and orangutans. The strategies used to analyse the data allowed to show that both great ape species expressed the four personality traits similarly across multiple distinct contexts and showed consistency in their behaviours.

As a result of a small dataset, in this project, some of the contexts (aggression, vigilance and locomotion) had to be excluded from the analyses. It would be interesting for future studies to extend the range of contexts to assess personality consistency. By following this approach, we could provide more insight into the diversity of expression of the different personality traits.

Future research would benefit from using the GT analysis to assess whether the personality trait or behaviour scores generalize across both contexts and time. Relying on the review of this theory (see Chapter 2; for further details regarding GT, see: Brennan, 2011; Shavelson, Webb, & Rowley, 1989), we were able to estimate separately the different sources of errors (context, time, interactions) that may have contributed to the variance in the personality assessment. With this theory, we were able to highlight inter-individual differences across the behaviours, determine how consistent the behaviours were across contexts, time as well as across both contexts and time. Furthermore, this statistical theory revealed variability in terms of both temporal and contextual consistency across subjects which provides further insight into individuality. Using such an approach to assess consistency of personality would greatly benefit the field as it provides a much more detailed perspective in terms of consistency on the personality traits targeted. Additionally, GT can be applied with a small sample size (Figueredo et al., 1995) and reliably measure welfare and well-being ratings (Robinson et al., 2017).

6.2.2. Alternative approach to measure personality

In the current thesis, we looked at both temporal and contextual consistency on the behavioural level as commonly done in nonhuman primate personality research (Koski, 2011b; Massen et al., 2013; Neumann et al., 2013; Seyfarth et al., 2012; Šlipogor et al., 2016; Tkaczynski et al., 2018). With this behaviour-focused approach, only behaviours that showed consistency across contexts and over time were considered for further analyses. The inconsistent behaviours, which are implicitly considered as weak indicators of a personality trait (Réale et al., 2007), were ignored in our analyses. However, it is

possible that those inconsistent behaviours could have provided extra information in terms of individuality.

Instead, using a broad list of behaviours, which are related conceptually to personality traits and based on empirical studies (conceptual approach: Uher, 2008b, 2008a; see Chapter 2 for a detailed review), rather than single behaviours would represent stronger indicators of the traits (Paunonen, 2001). The conceptual approach developed by Uher might be an alternative to consider in future animal personality studies as it is to some extent more inclusive than the approach used in this thesis. Generally speaking, such an approach could bring a different perspective about personality because including multiple behaviours would be likely to cover a wide range of facets of expression of a personality trait. To date, there is some empirical studies in both humans (Wu & Clark, 2003) and nonhuman primates (Tomassetti et al., 2019; Uher et al., 2013, 2008; Uher & Asendorpf, 2008; Uher & Visalberghi, 2016) research that measured personality consistency on the trait level and found evidence of contextual and temporal consistency. Yet, further investigation would be essential to enrich the current discussion in personality research.

Additionally, working on the conceptual level might also limit the risk of wrongly labelling a trait as it is sometimes done in animal personality research (jingle-jangle fallacies: see Chapter 2; Carter, Feeney, Marshall, Cowlishaw, & Heinsohn, 2013; Carter et al., 2012b). For instance, some researchers may define boldness based on behaviours that relate to risk-taking (Carter & Feeney, 2012; Wilson et al., 1994) whereas others might define boldness based on the expression of self-directed behaviours (Tkaczynski et al., 2018) or attention duration to a threatening playback (Neumann et al., 2013). However,

based on past research in primates, these behavioural indicators (i.e., self-directed behaviours or attention duration to a threatening playback) could be instead associated with anxiety (Aureli & Waal, 1997; Baker & Aureli, 1997; Kutsukake et al., 2012). It is important for future research to define explicitly the personality traits; the conceptual approach might help achieve this goal by considering multiple behaviours that reflect the personality trait targeted based on past empirical studies.

6.2.3. *Influence of the social environment*

Diverse social organisations characterise nonhuman primate society (Smuts et al., 2008) in which the behaviours of individuals may be affected by their conspecifics in many different ways during social interactions, such as affiliative (Jaeggi & Van Schaik, 2011; Mitani, 2009) and agonistic (Slocombe & Zuberbühler, 2005; Utami Atmoko et al., 2009). Depending on their needs, individuals may adjust their behaviours (Flack et al., 2004), so they could optimise their benefits in terms of fitness (Dingemanse & Wolf, 2013; Wolf & Weissing, 2012). Past research in nonhuman primates found some evidence of differences in the expression of personality traits across groups with different experience (Bard & Gardner, 1996; Koski, 2011b; Koski & Burkart, 2015) and personality adjustment from individuals immigrating into a new group (Sapolsky & Share, 2004).

Although no statistical analyses were conducted to compare the three orangutan groups of this study, the individuals seemed to show some plasticity in the expression of some of their personality-linked behaviours when measured across contexts. As mentioned earlier, however, we were not able to determine whether this variability of expression was influenced by specific members, age, or social environment differences. Building on this preliminary observations, and particularly, on the empirical studies that found evidence of

social influences on personality, it seems reasonable for future studies to investigate further to which extent personality can be shaped by the social environment (e.g., group composition, mother or friend's behaviours).

6.2.4. Evolutionary significance

Conducting research on such daily naturally occurring personality-linked behaviours in great apes helped contribute to some extent to a better understanding of the evolutionary history of human personality (Buss, 2014; Buss & Hawley, 2010; Buss & Penke, 2015). Of course, it is important to note that the aim of this thesis was not to compare the personality of the two great species; too many confounding factors would have had to be considered in the analyses such as differences in settings, demographics, or early experiences. Nonetheless, we showed that the naturalistic approach used to measure personality was applicable to both species. This naturalistic method relies on day-to-day behaviours (e.g., food sharing, gaze, self-touch) that naturally occur in ecologically-relevant contexts, such as play and feeding. Most of these behaviours and contexts are commonly expressed in various species, including humans.

Humans live in complex societies (Kaplan et al., 2009), where different social organizations may induce behavioural variability (Preuschoft & van Hooff, 1995), such as Westerners and Easterners who are, respectively, characterised as individualistic (independent and autonomous) and collectivist (group-orientated) cultures (Triandis, 2001). Regardless of the culture, the five personality dimensions that characterise humans' personality structure (Openness, Conscientiousness, Extraversion, Agreeableness, Neuroticism) have been reported to be universals (McCrae & Costa, 1997). Because these two cultures are characterised by the behaviours of people – e.g., the behaviours of

collectivist people are shaped by the norms instituted within the group – (Triandis, 2001; Triandis & Suh, 2002), we could expect to see different expression of the same personality traits across the cultures. For instance, collectivist people might show a higher level of sociability due to their group-oriented lifestyle compared to individualist people. In order to understand how humans developed such a diversity of expression in personality-linked behaviours, nonhuman primates, especially the great apes, seem to be an ideal model to study in the future.

Among nonhuman primates, there is a considerable variation in social group composition (Smuts et al., 2008). More specifically, among great ape species, the social structure varies from semi-solitary (orangutans) to highly social (chimpanzees) (McGrew, McGrew, Marchant, & Nishida, 1996). In the wild, orangutans and chimpanzees live in different social and ecological environments (Smuts et al., 2008); for instance, these two factors are likely to have had an impact on the evolution of their sociability trait where orangutans may show lower scores of sociability than chimpanzees.

It is important to note that besides phylogenetic, socioecological, past experience or demographic differences between the chimpanzees and orangutans of this research, both species showed commonalities in terms of personality consistency when measured using naturalistic observations. It is possible that this characteristic of consistency may be also present in the other two great species – bonobos and gorillas. It would be interesting to conduct such comparisons using the naturalistic method developed here, so we could further investigate this notion of consistency in other great ape species and possibly extend beyond these species by including other nonhuman primates. By doing so, we could gradually construct a map of personality expression, which could help perhaps identify the

different selective pressures that may have played a role in the emergence of personality in humans (Buss, 2014; Buss & Hawley, 2010; Nettle, 2006).

6.3 Conclusion

Despite facing some challenges in terms of collecting large dataset, the way used to approach the data allowed me to provide solid evidence of personality consistency, particularly contextual consistency, in sanctuary chimpanzees and rehabilitant orangutans using naturalistic observations. In this work, I provide a comprehensive view of the personality structure of two great ape species by focusing on behaviours that are related to sociability, boldness, exploratiVeness and anxiety. These personality-linked behaviours naturally occurred in multiple distinct contexts on multiple occasions. I demonstrate the importance of considering the context when examining nonhuman primate personality. Indeed, the context provides key information regarding the personality traits. Combining such a naturalistic approach with a different method of personality assessment (i.e., ratings or experimental observations) contributes to an improved understanding and a detailed description of personality traits. By doing so, we can better encompass the complexity of personality, in terms of how individuals express personality, how personality is measured or the factors that may influence its expression. Building upon this work, I suggest using a more conceptual approach when examining personality, so more behaviours reflecting the personality traits of interest are included which are likely to provide more variance across individuals. The naturalistic method developed in this thesis could favour future comparative work and could be advantageous for researchers who want to study the personality of free-ranging wildlife, such as the rehabilitant orangutans of this study.

References

Adams, M. J., King, J. E., & Weiss, A. (2012). The Majority of Genetic Variation in Orangutan Personality and Subjective Well-Being is Nonadditive. *Behavior Genetics*, *42*(4), 675–686. https://doi.org/10.1007/s10519-012-9537-y

Adams, M. J., Majolo, B., Ostner, J., Schülke, O., De Marco, A., Thierry, B., … Weiss, A. (2015). Personality structure and social style in macaques. *Journal of Personality and Social Psychology*, *109*, 338–353.

Allport, G. W. (1937). *Personality: A psychological interpretation*. Oxford, England: Holt.

Allport, G. W. (1961). *Pattern and growth in personality*. Oxford, England: Holt, Reinhart & Winston.

Allport, G. W., & Odbert, H. S. (1936). Trait names: A psycholexial study. *Psychological Monographs*, *47*(1).

Altmann, J. (1974). Observational study of behavior: Sampling methods. *Behaviour*, *49*(3–4), 227–267. https://doi.org/10.1163/156853974X00534

Altschul, D. M., Robinson, L. M., Coleman, K., Capitanio, J. P., & Wilson, V. A. D. (2019). An Exploration of the Relationships Among Facial Dimensions, Age, Sex, Dominance Status, and Personality in Rhesus Macaques (Macaca mulatta). *International Journal of Primatology*. https://doi.org/10.1007/s10764-019-00104-y

Altschul, D. M., Hopkins, W. D., Herrelko, E. S., Inoue-Murayama, M., Matsuzawa, T., King, J. E., … Weiss, A. (2018). Personality links with lifespan in chimpanzees. *ELife*, *7*. Retrieved from https://www.academia.edu/37564615/Personality_links_with_lifespan_in_chimpanzees

Amici, F., Call, J., Watzek, J., Brosnan, S., & Aureli, F. (2018). Social inhibition and behavioural flexibility when the context changes: A comparison across six primate species. *Scientific Reports*, *8*(1), 1–9. https://doi.org/10.1038/s41598-018-21496-6

Anestis, S. F. (2005). Behavioral style, dominance rank, and urinary cortisol in young chimpanzees (Pan troglodytes). *Behaviour*, *142*(9–10), 1245–1268. https://doi.org/10.1163/156853905774539418

Anestis, S. F. (2006). Testosterone in juvenile and adolescent male chimpanzees (Pan troglodytes): Effects of dominance rank, aggression, and behavioral style. *American Journal of Physical Anthropology*, *130*(4), 536–545. https://doi.org/10.1002/ajpa.20387

Araya-Ajoy, Y. G., & Dingemanse, N. J. (2017). Repeatability, heritability, and age-dependence of seasonal plasticity in aggressiveness in a wild passerine bird. *Journal of Animal Ecology*, 86(2), 227-238.

Archie, E. A., Tung, J., Clark, M., Altmann, J., & Alberts, S. C. (2014). Social affiliation matters: Both same-sex and opposite-sex relationships predict survival in wild female baboons. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1793), 20141261. https://doi.org/10.1098/rspb.2014.1261

Asendorpf, J. B., & Wilpers, S. (1998). Personality effects on social relationships. *Journal of Personality and Social Psychology*, *74*(6), 1531–1544. https://doi.org/10.1037/0022-3514.74.6.1531

Aureli, F., & Waal, F. B. M. D. (1997). Inhibition of social behavior in chimpanzees under high-density conditions. *American Journal of Primatology*, *41*(3), 213–228. https://doi.org/10.1002/(SICI)1098-2345(1997)41:3<213::AID-AJP4>3.0.CO;2-#

Baker, K. C., & Aureli, F. (1997). Behavioural Indicators of Anxiety: An Empirical Test in Chimpanzees. *Behaviour*, *134*(13), 1031–1050. https://doi.org/10.1163/156853997X00386

Baker, K. R., Lea, S. E. G., & Melfi, V. A. (2015). Comparative Personality Assessment of Three Captive Primate Species: Macaca nigra, Macaca sylvanus, and Saimiri sciureus. *International Journal of Primatology*, *36*(3), 625–646. https://doi.org/10.1007/s10764-015-9843-3

Bard, K. A., & Gardner, K. H. (1996). Influences on development in infant chimpanzees: Enculturation, temperament, and cognition. In A. Russon, K. A. Bard, & S. T. Parker (Eds.), *Reaching into thought: The minds of the great apes* (pp. 235–256). Cambridge; UK: Cambridge University Press.

Bardi, M., Shimizu, K., Fujita, S., Borgognini-Tarli, S., & Huffman, M. A. (2001). Hormonal Correlates of Maternal Style in Captive Macaques (Macaca fuscata and M. mulatta). *International Journal of Primatology*, *22*(4), 647–662. https://doi.org/10.1023/A:1010793719669

Barr, C. S., Newman, T. K., Lindell, S., Shannon, C., Champoux, M., Lesch, K. P., … Higley, J. D. (2004). Interaction Between Serotonin Transporter Gene Variation and RearingCondition in Alcohol Preference and Consumption in Female Primates. *Archives of General Psychiatry*, *61*(11), 1146–1152. https://doi.org/10.1001/archpsyc.61.11.1146

Barros, M., Boere, V., Mello, E. L., & Tomaz, C. (2002). Reactions to Potential Predators in Captive-Born Marmosets (Callithrix penicillata). *International Journal of Primatology*, *23*(2), 443–454. https://doi.org/10.1023/A:1013899931878

Bergvall, U. A., Schäpers, A., Kjellander, P., & Weiss, A. (2011). Personality and foraging decisions in fallow deer, Dama dama. *Animal Behaviour*, *81*(1), 101–112. https://doi.org/10.1016/j.anbehav.2010.09.018

Berry, J. W. (1997). Immigration, Acculturation, and Adaptation. *Applied Psychology*, *46*(1), 5–34. https://doi.org/10.1111/j.1464-0597.1997.tb01087.x

Berry, J. W. (2005). Acculturation: Living successfully in two cultures. *International Journal of Intercultural Relations*, *29*(6), 697–712. https://doi.org/10.1016/j.ijintrel.2005.07.013

Boesch, C., Boesch, P. C., & Boesch-Achermann, H. (2000). *The Chimpanzees of the Taï Forest: Behavioural Ecology and Evolution*. Oxford University Press.

Boissy, A. (1995). Fear and Fearfulness in Animals. *The Quarterly Review of Biology*, *70*(2), 165–191.

Bolig, R., Price, C. S., O'Neill, P. L., & Suomi, S. J. (1992). Subjective assessment of reactivity level and personality traits of rhesus monkeys. *International Journal of Primatology*, *13*(3), 287–306. https://doi.org/10.1007/BF02547817

Botero, M., MacDonald, S. E., & Miller, R. S. (2013). Anxiety-related behavior of orphan chimpanzees (Pan troglodytes schweinfurthii) at Gombe National Park, Tanzania. *Primates*, *54*(1), 21–26. https://doi.org/10.1007/s10329-012-0327-1

Box, H. O., Röhrhuber, B., & Smith, P. (1995). Female tamarins (Saguinus—Callitrichidae) feed more successfully than males in unfamiliar foraging tasks. *Behavioural Processes*, *34*(1), 3–11. https://doi.org/10.1016/0376-6357(94)00043-G

Bremner-Harrison, S., Prodohl, P. A., & Elwood, R. W. (2004). Behavioural trait assessment as a release criterion: Boldness predicts early death in a reintroduction programme of captive-bred swift fox (Vulpes velox). *Animal Conservation*, *7*(3), 313–320. https://doi.org/10.1017/S1367943004001490

Brennan, R. L. (1992). Generalizability Theory. *Educational Measurement: Issues and Practice*, *11*(4), 27–34. https://doi.org/10.1111/j.1745-3992.1992.tb00260.x

Brennan, R. L. (2001). *Generalizability Theory*. Retrieved from https://www.springer.com/gp/book/9780387952826

Brennan, R. L. (2003). Coefficients and Indices in Generalizability Theory. *Center for Advanced Studies in Measurement and Assessment, CASMA Research Report*, *1*, 1–44.

Brennan, R. L. (2011). Generalizability Theory and Classical Test Theory. *Applied Measurement in Education*, *24*(1), 1–21. https://doi.org/10.1080/08957347.2011.532417

Brent, L. J. N., Semple, S., MacLarnon, A., Ruiz-Lambides, A., Gonzalez-Martinez, J., & Platt, M. L. (2014). Personality Traits in Rhesus Macaques (Macaca mulatta) Are Heritable but Do Not Predict Reproductive Output. *International Journal of Primatology*, *35*(1), 188–209. https://doi.org/10.1007/s10764-013-9724-6

Bressan, P., & Zucchi, G. (2009). Human kin recognition is self- rather than family-referential. *Biology Letters*, *5*(3), 336–338. https://doi.org/10.1098/rsbl.2008.0789

Brosnan, S. F., Hopper, L. M., Richey, S., Freeman, H. D., Talbot, C. F., Gosling, S. D., … Schapiro, S. J. (2015). Personality influences responses to inequity and contrast in chimpanzees. *Animal Behaviour*, *101*, 75–87. https://doi.org/10.1016/j.anbehav.2014.12.019

Budaev, S. V. (2010). Using Principal Components and Factor Analysis in Animal Behaviour Research: Caveats and Guidelines. *Ethology*, *116*(5), 472–480. https://doi.org/10.1111/j.1439-0310.2010.01758.x

Budaev, S. V., Zworykin, D. D., & Mochek, A. D. (1999). Individual differences in parental care and behaviour profile in the convict cichlid: A correlation study. *Animal Behaviour*, *58*(1), 195–202. https://doi.org/10.1006/anbe.1999.1124

Buirski, P., Kellerman, H., Plutchik, R., Weininger, R., & Buirski, N. (1973). A field study of emotions, dominance, and social behavior in a group of baboons (Papio anubis). *Primates*, *14*(1), 67–78. https://doi.org/10.1007/BF01730516

Buirski, P., & Plutchik, R. (1991). Measurement of deviant behavior in a gombe chimpanzee: Relation to later behavior. *Primates*, *32*(2), 207–211. https://doi.org/10.1007/BF02381177

Buirski, P., Plutchik, R., & Kellerman, H. (1978). Sex differences, dominance, and personality in the chimpanzee. *Animal Behaviour*, *26*, 123–129. https://doi.org/10.1016/0003-3472(78)90011-8

Burger, J. M. (1997). Personality. Pacific Grove, Calif: Brooks/Cole Pub. Co

Burghardt, G. M., Bartmess-LeVasseur, J. N., Browning, S. A., Morrison, K. E., Stec, C. L., Zachau, C. E., & Freeberg, T. M. (2012). Perspectives – Minimizing Observer Bias in Behavioral Studies: A Review and Recommendations. *Ethology*, *118*(6), 511–517. https://doi.org/10.1111/j.1439-0310.2012.02040.x

Burkart, J. M., Strasser, A., & Foglia, M. (2009). Trade-offs between social learning and individual innovativeness in common marmosets, Callithrix jacchus. *Animal Behaviour*, *77*(5), 1291–1301. https://doi.org/10.1016/j.anbehav.2009.02.006

Buss, A. H. (1988). *Personality: Evolutionary heritage and human distinctiveness*. Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.

Buss, A. H. (2014). *Personality: Evolutionary Heritage and Human Distinctiveness*. Psychology Press.

Buss, D. M., & Hawley, P. H. (2010). *The Evolution of Personality and Individual Differences*. Oxford University Press.

Buss, D. M., & Penke, L. (2015). Evolutionary personality psychology. In *APA Handbooks in Psychology*. *APA handbook of personality and social psychology, Volume 4: Personality processes and individual differences* (pp. 3–29). https://doi.org/10.1037/14343-001

Byrne, G., & Suomi, S. J. (1995). Development of activity patterns, social interactions, and exploratory behavior in infant tufted capuchins (Cebus apella). *American Journal of Primatology*, *35*(4), 255–270. https://doi.org/10.1002/ajp.1350350402

Byrne, G., & Suomi, S. J. (1998). Relationship of early infant state measures to behavior over the first year of life in the tufted capuchin monkey (Cebus apella). *American Journal of Primatology*, *44*(1), 43–56. https://doi.org/10.1002/(SICI)1098-2345(1998)44:1<43::AID-AJP4>3.0.CO;2-X

Byrne, G., & Suomi, S. J. (2002). Cortisol reactivity and its relation to homecage behavior and personality ratings in tufted capuchin (Cebus apella) juveniles from birth to six years of age. *Psychoneuroendocrinology*, *27*(1), 139–154. https://doi.org/10.1016/S0306-4530(01)00041-5

Caine, N. G., Earle, H., & Reite, M. (1983). Personality traits of adolescent pig-tailed monkeys (Macaca nemestrina): An analysis of social rank and early separation experience. *American Journal of Primatology*, *4*(3), 253–260. https://doi.org/10.1002/ajp.1350040304

Capitanio, J. P. (1984). Early experience and social processes in rhesus macaques (Macaca mulatta): I. Dyadic social interaction. *Journal of Comparative Psychology*, *98*(1), 35–44. https://doi.org/10.1037/0735-7036.98.1.35

Capitanio, J. P. (1985). Early experience and social processes in rhesus macaques (Macaca mulatta): II. Complex social interaction. *Journal of Comparative Psychology*, *99*(2), 133–144. https://doi.org/10.1037/0735-7036.99.2.133

Capitanio, J. P. (1999). Personality dimensions in adult male rhesus macaques: Prediction of behaviors across time and situation. *American Journal of Primatology*, *47*(4), 299–320. https://doi.org/10.1002/(SICI)1098-2345(1999)47:4<299::AID-AJP3>3.0.CO;2-P

Capitanio, J. P. (2002). Sociability and responses to video playbacks in adult male rhesus monkeys (Macaca mulatta). *Primates*, *43*(3), 169–177. https://doi.org/10.1007/BF02629645

Capitanio, J. P., Mendoza, S. P., & Baroncelli, S. (1999). The Relationship of Personality Dimensions in Adult Male Rhesus Macaques to Progression of Simian Immunodeficiency Virus Disease. *Brain, Behavior, and Immunity*, *13*(2), 138–154. https://doi.org/10.1006/brbi.1998.0540

Capitanio, J. P., Mendoza, S. P., & Bentson, K. L. (2004). Personality characteristics and basal cortisol concentrations in adult male rhesus macaques (Macaca mulatta). *Psychoneuroendocrinology*, *29*(10), 1300–1308. https://doi.org/10.1016/j.psyneuen.2004.04.001

Carere, C., & Maestripieri, D. (Eds.). (2013). Animal personalities: behavior, physiology, and evolution. University of Chicago Press.

Carter, A. J., & Feeney, W. E. (2012). Taking a Comparative Approach: Analysing Personality as a Multivariate Behavioural Response across Species. *PLOS ONE*, *7*(7), e42440. https://doi.org/10.1371/journal.pone.0042440

Carter, A. J., Feeney, W. E., Marshall, H. H., Cowlishaw, G., & Heinsohn, R. (2013). Animal personality: What are behavioural ecologists measuring?: What are animal personality researchers measuring. *Biological Reviews*, *88*(2), 465–475. https://doi.org/10.1111/brv.12007

Carter, A., Goldizen, A., & Heinsohn, R. (2012). Personality and plasticity: temporal behavioural reaction norms in a lizard, the Namibian rock agama. *Animal Behaviour*, 84(2), 471-477

Carter, A. J., Marshall, H. H., Heinsohn, R., & Cowlishaw, G. (2012a). Evaluating animal personalities: Do observer assessments and experimental tests measure the same thing? *Behavioral Ecology and Sociobiology*, *66*(1), 153–160. https://doi.org/10.1007/s00265-011-1263-6

Carter, A. J., Marshall, H. H., Heinsohn, R., & Cowlishaw, G. (2012b). How not to measure boldness: Novel object and antipredator responses are not the same in wild baboons. *Animal Behaviour*, *84*(3), 603–609. https://doi.org/10.1016/j.anbehav.2012.06.015

Carter, A. J., Marshall, H. H., Heinsohn, R., & Cowlishaw, G. (2014). Personality predicts the propensity for social learning in a wild primate. *PeerJ*, *2*, e283. https://doi.org/10.7717/peerj.283

Castles, D. L., Whiten, A., & Aureli, F. (1999). Social anxiety, relationships and self-directed behaviour among wild female olive baboons. *Animal Behaviour*, *58*(6), 1207–1215. https://doi.org/10.1006/anbe.1999.1250

Cattell, R. B. (1943). The description of personality: Basic traits resolved into clusters. *The Journal of Abnormal and Social Psychology*, *38*(4), 476–506. https://doi.org/10.1037/h0054116

Chamove, A. S., Eysenck, H. J., & Harlow, H. F. (1972). Personality in Monkeys: Factor Analyses of Rhesus Social Behaviour. *Quarterly Journal of Experimental Psychology*, *24*(4), 496–504. https://doi.org/10.1080/14640747208400309

Chapman, B. B., Thain, H., Coughlin, J., & Hughes, W. O. H. (2011). Behavioural syndromes at multiple scales in Myrmica ants. *Animal Behaviour*, *82*, 391–397. https://doi.org/10.1016/j.anbehav.2011.05.019

Chen, R., Gu, Z., Wang, X., Sun, B., Xia, D.-P., & Li, J. (2018). Personality and its differences among adult free-ranging Tibetan macaques (Macaca thibetana) at Mt. Huangshan, China. *Acta Theriologica Sinica*, *38*(2), 117–127.

Clarke, A. S., & Lindburg, D. G. (1993). Behavioral contrasts between male cynomolgus and lion-tailed macaques. *American Journal of Primatology*, *29*(1), 49–59. https://doi.org/10.1002/ajp.1350290106

Clarke, A. S., & Snipes, M. (1998). Early behavioral development and temperamental traits in mother- vs peer-reared rhesus monkeys. *Primates*, *39*(4), 433–448. https://doi.org/10.1007/BF02557567

Clarke, A. S., & Boinski, S. (1995). Temperament in nonhuman primates. *American Journal of Primatology*, *37*(2), 103–125. https://doi.org/10.1002/ajp.1350370205

Clay, A. W., Bard, K. A., & Bloomsmith, M. A. (2018). Effects of sex and early rearing condition on adult behavior, health, and well-being in captive chimpanzees (Pan troglodytes). *Behavioural Processes*, *156*, 58–76. https://doi.org/10.1016/j.beproc.2017.06.011

Clay, A. W., Bloomsmith, M. A., Bard, K. A., Maple, T. L., & Marr, M. J. (2015). Long-term effects of infant attachment organization on adult behavior and health in nursery-reared, captive chimpanzees (Pan troglodytes). *Journal of Comparative Psychology*, *129*(2), 145–159. https://doi.org/10.1037/a0038901

Coleman, K., & Wilson, D. S. (1998). Shyness and boldness in pumpkinseed sunfish: Individual differences are context-specific. *Animal Behaviour*, *56*(4), 927–936. https://doi.org/10.1006/anbe.1998.0852

Cooper, G. M. (2000). Heredity, Genes, and DNA. *The Cell: A Molecular Approach. 2nd Edition*. Retrieved from https://www.ncbi.nlm.nih.gov/books/NBK9944/

Costa, P. T., & McCrae, R. R. (1988). Personality in adulthood: A six-year longitudinal study of self-reports and spouse ratings on the NEO Personality Inventory. *Journal of Personality and Social Psychology*, *54*(5), 853–863. https://doi.org/10.1037/0022-3514.54.5.853

Costa, P. T., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO PI-R) and NEO Five-Factor Inventory (NEO-FFI)*. Psychological Assessment Resources.

Crawford, M. P. (1938). A behavior rating scale for young chimpanzees. *Journal of Comparative Psychology*, *26*(1), 79–92. https://doi.org/10.1037/h0054503

Dahlbäck, O. (1990). Personality and risk-taking. *Personality and Individual Differences*, *11*(12), 1235–1242. https://doi.org/10.1016/0191-8869(90)90150-P

Dahlbäck, O. (1991). Accident-proneness and risk-taking. *Personality and Individual Differences*, *12*(1), 79–85. https://doi.org/10.1016/0191-8869(91)90134-W

Damerius, L. A., Forss, S. I. F., Kosonen, Z. K., Willems, E. P., Burkart, J. M., Call, J., … van Schaik, C. P. (2017). Orientation toward humans predicts cognitive performance in orang-utans. *Scientific Reports*, *7*, 40052. https://doi.org/10.1038/srep40052

Damerius, L. A., Graber, S. M., Willems, E. P., & van Schaik, C. P. (2017). Curiosity boosts orang-utan problem-solving ability. *Animal Behaviour*, *134*, 57–70. https://doi.org/10.1016/j.anbehav.2017.10.005

Dammhahn, M. (2012). Are personality differences in a small iteroparous mammal maintained by a life-history trade-off? | Proceedings of the Royal Society of London B: Biological Sciences. *Proceedings of The Royal Society B*, *13*, 2645–2651. https://doi.org/10.1098/rspb.2012.0212

Dammhahn, M., & Almeling, L. (2012). Is risk taking during foraging a personality trait? A field test for cross-context consistency in boldness. *Animal Behaviour*, *84*, 1131–1139.

Davidson, R. J., Kalin, N. H., & Shelton, S. (1993). Lateralized response to diazepam predicts temperamental style in rhesus monkeys. *Behavioral Neuroscience*, *107*(6), 1106–1110. https://doi.org/10.1037/0735-7044.107.6.1106

Davila-Ross, M., Allcock, B., Thomas, C., & Bard, K. A. (2011). Aping expressions? Chimpanzees produce distinct laugh types when responding to laughter of others. *Emotion*, *11*(5), 1013–1020. https://doi.org/10.1037/a0022594

Delgado, R. A., & van Schaik, C. P. (2000). The behavioral ecology and conservation of the orangutan (Pongo pygmaeus): A tale of two islands. *Evolutionary Anthropology: Issues, News, and Reviews*, *9*(5), 201–218. https://doi.org/10.1002/1520-6505(2000)9:5<201::AID-EVAN2>3.0.CO;2-Y

Digman, J. m. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology*, *41*(1), 417. https://doi.org/10.1146/annurev.ps.41.020190.002221

Dingemanse, N. J., Both, C., Drent, P. J., & Tinbergen, J. M. (2004). Fitness consequences of avian personalities in a fluctuating environment. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *271*(1541), 847–852. https://doi.org/10.1098/rspb.2004.2680

Dingemanse, N. J., Bouwman, K. M., Van De Pol, M., van Overveld, T., Patrick, S. C., Matthysen, E., & Quinn, J. L. (2012). Variation in personality and behavioural plasticity across four populations of the great tit Parus major. *Journal of Animal Ecology*, 81(1), 116-126

Dingemanse, N. J., Kazem, A. J., Réale, D., & Wright, J. (2010). Behavioural reaction norms: animal personality meets individual plasticity. *Trends in ecology & evolution*, 25(2), 81-89

Dingemanse, N. J., & Réale, D. (2005). Natural selection and animal personality. *Behaviour*, *142*(9–10), 1159–1184. https://doi.org/10.1163/156853905774539445

Dingemanse, N. J., & Wolf, M. (2013). Between-individual differences in behavioural plasticity within populations: Causes and consequences. *Animal Behaviour*, *85*(5), 1031–1039. https://doi.org/10.1016/j.anbehav.2012.12.032

Dugatkin, L. A. (1992). Tendency to inspect predators predicts mortality risk in the guppy (Poecilia reticulata). *Behavioral Ecology*, *3*(2), 124–127. https://doi.org/10.1093/beheco/3.2.124

Dugatkin, L. A. (2013). The Evolution of Risk-Taking. *Cerebrum: The Dana Forum on Brain Science*, *2013*. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3600861/

Dutton, D. M. (2008). Subjective assessment of chimpanzee (Pan troglodytes) personality: Reliability and stability of trait ratings. *Primates; Journal of Primatology*, *49*(4), 253–259. https://doi.org/10.1007/s10329-008-0094-1

Dutton, D. M., Clark, R. A., & Dickins, D. W. (1997). Personality in Captive Chimpanzees: Use of a Novel Rating Procedure. *International Journal of Primatology*, *18*(4), 539–552. https://doi.org/10.1023/A:1026311222491

Eaton, L. G., & Funder, D. C. (2001). Emotional experience in daily life: Valence, variability, and rate of change. *Emotion*, *1*(4), 413–421. https://doi.org/10.1037/1528-3542.1.4.413

Ebenau, A., von Borell, C., Penke, L., Ostner, J., & Schülke, O. (2019). Personality homophily affects male social bonding in wild Assamese macaques, Macaca assamensis. *Animal Behaviour*, *155*, 21–35. https://doi.org/10.1016/j.anbehav.2019.05.020

Eckardt, W., Steklis, H. D., Steklis, N. G., Fletcher, A. W., Stoinski, T. S., & Weiss, A. (2015). Personality dimensions and their behavioral correlates in wild Virunga mountain gorillas (Gorilla beringei beringei). *Journal of Comparative Psychology (Washington, D.C.: 1983)*, *129*(1), 26–41. https://doi.org/10.1037/a0038370

Elgar, M. A. (1989). Predator Vigilance and Group Size in Mammals and Birds: A Critical Review of the Empirical Evidence. *Biological Reviews*, *64*(1), 13–33. https://doi.org/10.1111/j.1469-185X.1989.tb00636.x

Eysenck, H. J. (1952). *The scientific study of personality*. Oxford, England: Macmillan.

Eysenck, H. J. (2013). *The structure of human personality (Psychology Revivals)*. Routledge.

Fagen, R. (1981). *Animal play behavior*. Oxford University Press.

Fairbanks, L. A. (1993). Risk-Taking By Juvenile Vervet Monkeys. *Behaviour*, *124*(1–2), 57–72. https://doi.org/10.1163/156853993X00506

Fairbanks, L. A. (2001). Individual differences in response to a stranger: Social impulsivity as a dimension of temperament in vervet monkeys (Cercopithecus aethiops sabaeus). *Journal of Comparative Psychology*, *115*(1), 22–28. https://doi.org/10.1037/0735-7036.115.1.22

Fairbanks, L. A., Bailey, J. N., Breidenthal, S. E., Laudenslager, M. L., Kaplan, J. R., & Jorgensen, M. J. (2011). Environmental stress alters genetic regulation of novelty seeking in vervet monkeys. *Genes, Brain and Behavior*, *10*(6), 683–688. https://doi.org/10.1111/j.1601-183X.2011.00707.x

Fairbanks, L. A., Fontenot, M. B., Phillips-Conroy, J. E., Jolly, C. J., Kaplan, J. R., & Mann, J. J. (1999). CSF Monoamines, Age and Impulsivity in Wild Grivet Monkeys (Cercopithecus aethiops aethiops). *Brain, Behavior and Evolution*, *53*(5–6), 305–312. https://doi.org/10.1159/000006601

Fairbanks, L. A., Jorgensen, M. J., Huff, A., Blau, K., Hung, Y.-Y., & Mann, J. J. (2004). Adolescent impulsivity predicts adult dominance attainment in male vervet monkeys. *American Journal of Primatology*, *64*(1), 1–17. https://doi.org/10.1002/ajp.20057

Fairbanks, L. A., & McGuire, M. T. (1993). Maternal protectiveness and response to the unfamiliar in vervet monkeys. *American Journal of Primatology*, *30*(2), 119–129. https://doi.org/10.1002/ajp.1350300204

Fairbanks, L. A., Melega, W. P., Jorgensen, M. J., Kaplan, J. R., & McGuire, M. T. (2001). Social Impulsivity Inversely Associated with CSF 5-HIAA and Fluoxetine Exposure in Vervet Monkeys. *Neuropsychopharmacology*, *24*(4), 370–378. https://doi.org/10.1016/S0893-133X(00)00211-6

Fairbanks, L. A., Newman, T. K., Bailey, J. N., Jorgensen, M. J., Breidenthal, S. E., Ophoff, R. A., … Rogers, J. (2004). Genetic contributions to social impulsivity and aggressiveness in vervet monkeys. *Biological Psychiatry*, *55*(6), 642–647. https://doi.org/10.1016/j.biopsych.2003.12.005

Faughn, C., Marrus, N., Shuman, J., Ross, S. R., Constantino, J. N., Pruett, J. R., & Povinelli, D. J. (2015). Brief Report: Chimpanzee Social Responsiveness Scale (CSRS) Detects Individual Variation in Social Responsiveness for Captive Chimpanzees. *Journal of Autism and Developmental Disorders*, *45*(5), 1483–1488. https://doi.org/10.1007/s10803-014-2273-9

Ferreira, R. G., Mendl, M., Wagner, P. G. C., Araujo, T., Nunes, D., & Mafra, A. L. (2016). Coping strategies in captive capuchin monkeys (Sapajus spp.). *Applied Animal Behaviour Science*, *176*, 120–127. https://doi.org/10.1016/j.applanim.2015.12.007

Field, A. (2013). *Discovering Statistics Using IBM SPSS Statistics* (4th ed.). Sage Publications Ltd.

Figueredo, A. J., Cox, R. L., & Rhine, R. J. (1995). A Generalizability Analysis of Subjective Personality Assessments in the Stumptail Macaque and the Zebra Finch. *Multivariate Behavioral Research*, *30*(2), 167–197. https://doi.org/10.1207/s15327906mbr3002_3

Flack, J. C., Jeannotte, L. A., & de Waal, F. B. M. (2004). Play signaling and the perception of social rules by juvenile chimpanzees (Pan troglodytes). *Journal of Comparative Psychology (Washington, D.C.: 1983)*, *118*(2), 149–159. https://doi.org/10.1037/0735-7036.118.2.149

Foerster, S., Franz, M., Murray, C. M., Gilby, I. C., Feldblum, J. T., Walker, K. K., & Pusey, A. E. (2016). Chimpanzee females queue but males compete for social status. *Scientific Reports*, *6*, 35404. https://doi.org/10.1038/srep35404

Forss, S. I. F., Schuppli, C., Haiden, D., Zweifel, N., & van Schaik, C. P. (2015). Contrasting responses to novelty by wild and captive orangutans: Novelty Response in Orangutans. *American Journal of Primatology*, *77*(10), 1109–1121. https://doi.org/10.1002/ajp.22445

Forss, S. I. F., Willems, E., Call, J., & van Schaik, C. P. (2016). *Cognitive differences between orang-utan species: A test of the cultural intelligence hypothesis*. https://doi.org/10.1038/srep30516

Freedman, L. Z., & Rosvold, H. E. (1962). Sexual, aggressive and anxious behavior in the laboratory macaque. *Journal of Nervous and Mental Disease*, *134*(1), 18–27. https://doi.org/10.1097/00005053-196213410-00002

Freeman, H., Brosnan, S. F., Hopper, L. M., Lambeth, S. P., Schapiro, S. J., & Gosling, S. D. (2013). Developing a Comprehensive and Comparative Questionnaire for Measuring Personality in Chimpanzees Using a Simultaneous Top-Down/Bottom-Up Design. *American Journal of Primatology*, *75*, 1042–1053.

Freeman, H., & Gosling, S. D. (2010). Personality in nonhuman primates: A review and evaluation of past research. *American Journal of Primatology*, *72*(8), 653–671. https://doi.org/10.1002/ajp.20833

Freeman, H., Gosling, S. D., & Schapiro, S. J. (2011). Comparison of methods for assessing personality in nonhuman primates. In A. Weiss, J. E. King, & L. E. Murray, *Personality and temperament in nonhuman primates* (pp. 17–40). https://doi.org/10.1007/978-1-4614-0176-6_2

French, J. A. (1981). Individual differences in play inMacaca fuscata: The role of maternal status and proximity. *International Journal of Primatology*, *2*(3), 237–246. https://doi.org/10.1007/BF02739332

Funder, D. C. (2001). Personality. *Annual Review of Psychology*, *52*(1), 197–221. https://doi.org/10.1146/annurev.psych.52.1.197

Funder, D. C. (2006). Towards a resolution of the personality triad: Persons, situations, and behaviors. *Journal of Research in Personality*, *40*(1), 21–34. https://doi.org/10.1016/j.jrp.2005.08.003

Funder, D. C., & Colvin, R. (1991). Explorations in behavioral consistency: Properties of persons, situations, and behaviors. *Journal of Personality and Social Psychology*, *60*(5), 773–794. https://doi.org/10.1037/0022-3514.60.5.773

Furr, R. M. (2009). Personality psychology as a truly behavioural science. *European Journal of Personality*, *23*(5), 369–401. https://doi.org/10.1002/per.724

Furr, R. M., & Funder, D. C. (2004). Situational similarity and behavioral consistency: Subjective, objective, variable-centered, and person-centered approaches. *Journal of Research in Personality*, *38*(5), 421–447. https://doi.org/10.1016/j.jrp.2003.10.001

Galdikas, B. M. F. (1985). Adult Male Sociality and Reproductive Tactics among Orangutans at Tanjung Puting. *Folia Primatologica*, *45*(1), 9–24. https://doi.org/10.1159/000156188

Galdikas, B. M. F. (1995). Social and Reproductive Behavior of Wild Adolescent Female Orangutans. In R. D. Nadler, B. F. M. Galdikas, L. K. Sheeran, & N. Rosen (Eds.), *The Neglected Ape* (pp. 163–182). https://doi.org/10.1007/978-1-4899-1091-2_17

Garai, C., Weiss, A., Arnaud, C., & Furuichi, T. (2016). Personality in wild bonobos (Pan paniscus). *American Journal of Primatology*, *78*(11), 1178–1189. https://doi.org/10.1002/ajp.22573

Gold, K. C., & Maple, T. L. (1994). Personality assessment in the gorilla and its utility as a management tool. *Zoo Biology*, *13*(5), 509–522. https://doi.org/10.1002/zoo.1430130513

Goldberg, L. R. (1990). An Alternative 'Description of Personality': The Big-Five Factor Structure. *Journal of Personality & Social Psychology*, *59*, 1216–1229.

Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American Psychologist*, *48*(1), 26–34. https://doi.org/10.1037/0003-066X.48.1.26

Goodall, J. (1986). *The chimpanzees of Gombe: Patterns of behavior*. Cambridge, Mass: Belknap Press of Harvard University Press.

Goodman, M., Porter, C. A., Czelusniak, J., Page, S. L., Schneider, H., Shoshani, J., … Groves, C. P. (1998). Toward a Phylogenetic Classification of Primates Based on DNA Evidence Complemented by Fossil Evidence. *Molecular Phylogenetics and Evolution*, *9*(3), 585–598. https://doi.org/10.1006/mpev.1998.0495

Gosling, S. D. (2001). From mice to men: What can we learn about personality from animal research? *Psychological Bulletin*, *127*(1), 45–86. https://doi.org/10.1037/0033-2909.127.1.45

Gosling, S. D. (2008). Personality in Non-human Animals -. *Social and Personality Psychology Compass*, *2*(2), 985–1001.

Gosling, S. D., & John, O. P. (1999). Personality Dimensions in non human animals: A cross-species review. *Psychological Science*.

Gray, J. A., & MacNaughton, N. (2003). *The neuropsychology of anxiety: An enquiry into the functions of the septo-hippocampal system* (2nd ed). Oxford: Oxford University Press.

Gruber, T., Luncz, L., Mörchen, J., Schuppli, C., Kendal, R. L., & Hockings, K. (2019). Cultural change in animals: A flexible behavioural adaptation to human disturbance. *Palgrave Communications*, *5*(1), 1–9. https://doi.org/10.1057/s41599-019-0271-4

Hankerson, S. J., & Caine, N. G. (2004). Pre-retirement predator encounters alter the morning behavior of captive marmosets (Callithrix geoffroyi). *American Journal of Primatology*, *63*(2), 75–85. https://doi.org/10.1002/ajp.20040

Harlow, R. E., & Cantor, N. (1995). To whom do people turn when things go poorly? Task orientation and functional social contacts. *Journal of Personality and Social Psychology*, *69*(2), 329–340. https://doi.org/10.1037/0022-3514.69.2.329

Haslerud, G. M. (1938). The effect of movement of stimulus objects upon avoidance reactions in chimpanzees. *Journal of Comparative Psychology*, *25*(3), 507–528. https://doi.org/10.1037/h0063562

Heath-Lange, S., Ha, J. C., & Sackett, G. P. (1999). Behavioral measurement of temperament in male nursery-raised infant macaques and baboons. *American Journal of Primatology*, *47*(1), 43–50. https://doi.org/10.1002/(SICI)1098-2345(1999)47:1<43::AID-AJP5>3.0.CO;2-M

Hebb, D. O. (1949). Temperament in chimpanzees: I. Method of analysis. *Journal of Comparative and Physiological Psychology*, *42*(3), 192–206. https://doi.org/10.1037/h0056842

Hernández-Lloreda, M. V., & Colmenares, F. (2006). The utility of generalizability theory in the study of animal behaviour. *Animal Behaviour*, *71*(4), 983–988. https://doi.org/10.1016/j.anbehav.2005.04.023

Herrelko, E. S., Vick, S.-J., & Buchanan-Smith, H. M. (2012). Cognitive Research in Zoo-Housed Chimpanzees: Influence of Personality and Impact on Welfare. *American Journal of Primatology*, *74*(9), 828–840. https://doi.org/10.1002/ajp.22036

Herrmann, E., Hare, B., Cissewski, J., & Tomasello, M. (2011). A comparison of temperament in nonhuman apes and human infants. *Developmental Science*, *14*(6), 1393–1405. https://doi.org/10.1111/j.1467-7687.2011.01082.x

Hewes, M. E., & Chaves-Campos, J. (2018). Boldness related to size in the hermit crab Coenobita compressus at undisturbed, but not disturbed beach. *Ethology*, *124*(8), 570–578. https://doi.org/10.1111/eth.12766

Hinde, K., Skibiel, A. L., Foster, A. B., Del Rosso, L., Mendoza, S. P., & Capitanio, J. P. (2015). Cortisol in mother's milk across lactation reflects maternal life history and predicts infant temperament. *Behavioral Ecology*, *26*(1), 269–281. https://doi.org/10.1093/beheco/aru186

Hobolth, A., Dutheil, J. Y., Hawks, J., Schierup, M. H., & Mailund, T. (2011). Incomplete lineage sorting patterns among human, chimpanzee, and orangutan suggest recent orangutan speciation and widespread selection. *Genome Research*, *21*(3), 349–356. https://doi.org/10.1101/gr.114751.110

Holm, S. (1979). A Simple Sequentially Rejective Multiple Test Procedure. *Scandinavian Journal of Statistics*, *6*(2), 65–70. Retrieved from JSTOR.

Hopkins, W. D., Donaldson, Z. R., & Young, L. J. (2012). A polymorphic indel containing the RS3 microsatellite in the 5′ flanking region of the vasopressin V1a receptor gene is associated with chimpanzee (Pan troglodytes) personality. *Genes, Brain and Behavior*, *11*(5), 552–558. https://doi.org/10.1111/j.1601-183X.2012.00799.x

Hopper, L. M., Price, S. A., Freeman, H. D., Lambeth, S. P., Schapiro, S. J., & Kendal, R. L. (2014). Influence of personality, age, sex, and estrous state on chimpanzee problem-solving success. *Animal Cognition*, *17*(4), 835–847. https://doi.org/10.1007/s10071-013-0715-y

Humle, T., Maisels, F., Oates, J. F., Plumptre, A., & Williamson, E. A. (2016). *Pan troglodytes* [Data set]. https://doi.org/10.2305/IUCN.UK.2016-2.RLTS.T15933A17964454.en

Ickes, W., Snyder, M., & Garcia, S. (1997). Personality influences on the choice of situations. In *R. Hogan, J. Johnson, & S. Briggs (Eds.), Handbook of personality psychology* (pp. 165–195). New York: Academic Press.

IJzendoorn, M. H. van, Bard, K. A., Bakermans-Kranenburg, M. J., & Ivan, K. (2009). Enhancement of attachment and cognitive development of young nursery-reared chimpanzees in responsive versus standard care. *Developmental Psychobiology*, *51*(2), 173–185. https://doi.org/10.1002/dev.20356

Inoue-Murayama, M., Yokoyama, C., Yamanashi, Y., & Weiss, A. (2018). Common marmoset (Callithrix jacchus) personality, subjective well-being, hair cortisol level and AVPR1a , OPRM1 , and DAT genotypes. *Scientific Reports*, *8*(1), 10255. https://doi.org/10.1038/s41598-018-28112-7

Itoh, K. (2002). Personality research with non-human primates: Theoretical formulation and methods. *Primates*, *43*(3), 249–261. https://doi.org/10.1007/BF02629652

Iwanicki, S., & Lehmann, J. (2015). Behavioral and trait rating assessments of personality in common marmosets (Callithrix jacchus). *Journal of Comparative Psychology*, *129*(3), 205–217. https://doi.org/10.1037/a0039318

Jaeggi, A. V., Dunkel, L. P., Noordwijk, M. A. V., Wich, S. A., Sura, A. A. L., & Schaik, C. P. V. (2010). Social learning of diet and foraging skills by wild immature Bornean orangutans: Implications for culture. *American Journal of Primatology*, *72*(1), 62–71. https://doi.org/10.1002/ajp.20752

Jaeggi, A. V., & Van Schaik, C. P. (2011). The evolution of food sharing in primates. *Behavioral Ecology and Sociobiology*, *65*(11), 2125. https://doi.org/10.1007/s00265-011-1221-3

James, A. S., Groman, S. M., Seu, E., Jorgensen, M., Fairbanks, L. A., & Jentsch, J. D. (2007). Dimensions of Impulsivity Are Associated with Poor Spatial Working Memory Performance in Monkeys. *Journal of Neuroscience*, *27*(52), 14358–14364. https://doi.org/10.1523/JNEUROSCI.4508-07.2007

Jarrett, J. D., Bonnell, T. R., Young, C., Barrett, L., & Henzi, S. P. (2018). Network integration and limits to social inheritance in vervet monkeys. *Proceedings of the Royal Society B: Biological Sciences*, *285*(1876), 20172668. https://doi.org/10.1098/rspb.2017.2668

John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative Big Five trait taxonomy: History, measurement, and conceptual issues. In *Handbook of personality: Theory and research, 3rd ed* (pp. 114–158). New York, NY, US: The Guilford Press.

John, O. P., Robins, R. W., & Pervin, L. A. (2008). *Handbook of Personality, Third Edition: Theory and Research*. Guilford Press.

John, O. P., & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality theory and research* (pp. 102–138). New York, US: Guilford Press.

Johnson, Z., Brent, L., Alvarenga, J. C., Comuzzie, A. G., Shelledy, W., Ramirez, S., … Rogers, J. (2015). Genetic Influences on Response to Novel Objects and Dimensions of Personality in Papio Baboons. *Behavior Genetics*, *45*(2), 215–227. https://doi.org/10.1007/s10519-014-9702-6

Josephs, N., Bonnell, T., Dostie, M., Barrett, L., & Henzi, S. P. (2016). Working the crowd: Sociable vervets benefit by reducing exposure to risk. *Behavioral Ecology*, *27*(4), 988–994. https://doi.org/10.1093/beheco/arw003

Judge, T. A., & Bono, J. E. (2000). Five-factor model of personality and transformational leadership. *Journal of Applied Psychology*, *85*(5), 751–765. https://doi.org/10.1037//0021-9010.85.5.751

Kalin, N. H., & Shelton, S. E. (2003). Nonhuman Primate Models to Study Anxiety, Emotion Regulation, and Psychopathology. *Annals of the New York Academy of Sciences*, *1008*(1), 189–200. https://doi.org/10.1196/annals.1301.021

Kalin, N. H., Shelton, S. E., & Davidson, R. J. (2007). Role of the Primate Orbitofrontal Cortex in Mediating Anxious Temperament. *Biological Psychiatry*, *62*(10), 1134–1139. https://doi.org/10.1016/j.biopsych.2007.04.004

Kalin, N. H., Shelton, S. E., Davidson, R. J., & Kelley, A. E. (2001). The Primate Amygdala Mediates Acute Fear But Not the Behavioral and Physiological Components of Anxious Temperament. *Journal of Neuroscience*, *21*(6), 2067–2074. https://doi.org/10.1523/JNEUROSCI.21-06-02067.2001

Kalin, N. H., Shelton, S. E., Fox, A. S., Oakes, T. R., & Davidson, R. J. (2005). Brain Regions Associated with the Expression and Contextual Regulation of Anxiety in Primates. *Biological Psychiatry*, *58*(10), 796–804. https://doi.org/10.1016/j.biopsych.2005.05.021

Kalin, N. H., Shelton, S. E., Rickman, M., & Davidson, R. J. (1998). Individual differences in freezing and cortisol in infant and mother rhesus monkeys. *Behavioral Neuroscience*, *112*(1), 251–254. https://doi.org/10.1037/0735-7044.112.1.251

Kaplan, H. S., Hooper, P. L., & Gurven, M. (2009). The evolutionary and ecological roots of human social organization. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1533), 3289–3299. https://doi.org/10.1098/rstb.2009.0115

Kaplan, J. R., Manuck, S. B., Fontenot, M. B., & Mann, J. J. (2002). Central Nervous System Monoamine Correlates of Social Dominance in Cynomolgus Monkeys (Macaca fascicularis). *Neuropsychopharmacology*, *26*(4), 431–443. https://doi.org/10.1016/S0893-133X(01)00344-X

Kaufman, A. B., & Rosenthal, R. (2009). Can you believe my eyes? The importance of interobserver reliability statistics in observations of animal behaviour. *Animal Behaviour*, *78*(6), 1487–1491. https://doi.org/10.1016/j.anbehav.2009.09.014

Kenrick, D. t. (1), & Funder, D. c. (2). (1988). Profiting From Controversy: Lessons From the Person-Situation Debate. *American Psychologist*, *43*(1), 23–34. https://doi.org/10.1037/0003-066X.43.1.23

Kensinger, E. A., & Schacter, D. L. (2006). Reality monitoring and memory distortion: Effects of negative, arousing content. *Memory & Cognition*, *34*(2), 251–260. https://doi.org/10.3758/BF03193403

King, J. E., & Figueredo, A. J. (1997). The Five-Factor Model plus Dominance in Chimpanzee Personality. *Journal of Research in Personality*, *31*(2), 257–271. https://doi.org/10.1006/jrpe.1997.2179

King, J. E., & Landau, V. I. (2003). Can chimpanzee (Pan troglodytes) happiness be estimated by human raters? *Journal of Research in Personality*, *37*(1). Retrieved from http://ac.els-cdn.com/S0092656602005275/1-s2.0-S0092656602005275-main.pdf?_tid=cc6f6d8c-0835-11e5-9e2c-00000aacb35f&acdnat=1433146466_244ef00747c761acf81a14849b5cc02d

King, J. E., Weiss, A., & Farmer, K. H. (2005). A Chimpanzee (Pan troglodytes) Analogue of Cross-National Generalization of Personality Structure: Zoological Parks and an African Sanctuary. *Journal of Personality*, *73*(2), 389–410. https://doi.org/10.1111/j.1467-6494.2005.00313.x

King, J. E., Weiss, A., & Sisco, M. M. (2008). Aping humans: Age and sex effects in chimpanzee (Pan troglodytes) and human (Homo sapiens) personality. *Journal of Comparative Psychology*, *122*(4), 418–427. https://doi.org/10.1037/a0013125

Konečná, M., Lhota, S., Weiss, A., Urbánek, T., Adamová, T., & Pluháček, J. (2008). Personality in free-ranging Hanuman langur (Semnopithecus entellus) males: Subjective ratings and recorded behavior. *Journal of Comparative Psychology*, *122*(4), 379–389. https://doi.org/10.1037/a0012625

Konečná, M., Weiss, A., Lhota, S., & Wallner, B. (2012). Personality in Barbary macaques (Macaca sylvanus): Temporal stability and social rank. *Journal of Research in Personality*, *46*(5), 581–590. https://doi.org/10.1016/j.jrp.2012.06.004

Koolhaas, J. M., Korte, S. M., De Boer, S. F., Van Der Vegt, B. J., Van Reenen, C. G., Hopster, H., … Blokhuis, H. J. (1999). Coping styles in animals: Current status in behavior and stress-physiology. *Neuroscience and Biobehavioral Reviews*, *23*(7), 925–935.

Koppel, J., Brown, A. D., Stone, C. B., Coman, A., & Hirst, W. (2013). Remembering President Barack Obama's inauguration and the landing of US Airways Flight 1549: A comparison of the predictors of autobiographical and event memory. *Memory*, *21*(7), 798–806. https://doi.org/10.1080/09658211.2012.756040

Koski, S. E. (2011a). How to Measure Animal Personality and Why Does It Matter? Integrating the Psychological and Biological Approaches to Animal Personality. In M. Inoue-Murayama, S. Kawamura, & A. Weiss (Eds.), *From Genes to Animal Behavior* (pp. 115–136). Retrieved from http://link.springer.com/10.1007/978-4-431-53892-9_5

Koski, S. E. (2011b). Social personality traits in chimpanzees: Temporal stability and structure of behaviourally assessed personality traits in three captive populations. *Behavioral Ecology and Sociobiology*, *65*(11), 2161–2174. https://doi.org/10.1007/s00265-011-1224-0

Koski, S. E. (2014). Broader horizons for animal personality research. *Behavioral and Evolutionary Ecology*, *2*, 70. https://doi.org/10.3389/fevo.2014.00070

Koski, S. E., Buchanan-Smith, H. M., Ash, H., Burkart, J. M., Bugnyar, T., & Weiss, A. (2017). Common marmoset (Callithrix jacchus) personality. *Journal of Comparative Psychology*, *131*(4), 326–336. https://doi.org/10.1037/com0000089

Koski, S. E., & Burkart, J. M. (2015). Common marmosets show social plasticity and group-level similarity in personality. *Scientific Reports*, *5*, 8878.

Kuhar, C. W., Stoinski, T. S., Lukas, K. E., & Maple, T. L. (2006). Gorilla Behavior Index revisited: Age, housing and behavior. *Applied Animal Behaviour Science*, *96*(3), 315–326. https://doi.org/10.1016/j.applanim.2005.06.004

Kutsukake, N. (2003). Assessing Relationship Quality and Social Anxiety Among Wild Chimpanzees Using Self-Directed Behaviour. *Behaviour*, *140*(8), 1153–1171. https://doi.org/10.1163/156853903322589687

Kutsukake, N. (2006). The Context and Quality of Social Relationships Affect Vigilance Behaviour in Wild Chimpanzees. *Ethology*, *112*(6), 581–591. https://doi.org/10.1111/j.1439-0310.2006.01200.x

Kutsukake, N., Teramoto, M., Homma, S., Mori, Y., Matsudaira, K., Kobayashi, H., … Hasegawa, T. (2012). Individual Variation in Behavioural Reactions to Unfamiliar Conspecific Vocalisation and Hormonal Underpinnings in Male Chimpanzees. *Ethology*, *118*(3), 269–280. https://doi.org/10.1111/j.1439-0310.2011.02009.x

Kuze, N., Malim, T. P., & Kohshima, S. (2005). Developmental changes in the facial morphology of the Borneo orangutan (Pongo pygmaeus): Possible signals in visual communication. *American Journal of Primatology*, *65*(4), 353–376. https://doi.org/10.1002/ajp.20121

Latzman, R. D., Freeman, H. D., Schapiro, S. J., & Hopkins, W. D. (2015). The contribution of genetics and early rearing experiences to hierarchical personality dimensions in chimpanzees (Pan troglodytes). *Journal of Personality and Social Psychology*, *109*(5), 889–900. https://doi.org/10.1037/pspp0000040

Latzman, R. D., Hecht, L. K., Freeman, H. D., Schapiro, S. J., & Hopkins, W. D. (2015). Neuroanatomical correlates of personality in chimpanzees (Pan troglodytes): Associations between personality and frontal cortex. *NeuroImage*, *123*, 63–71. https://doi.org/10.1016/j.neuroimage.2015.08.041

Latzman, R. D., Hopkins, W. D., Keebaugh, A. C., & Young, L. J. (2014). Personality in Chimpanzees (Pan troglodytes): Exploring the Hierarchical Structure and Associations with the Vasopressin V1A Receptor Gene. *PLOS ONE*, *9*(4), e95741. https://doi.org/10.1371/journal.pone.0095741

Laudenslager, M. L., & Boccia, M. L. (1996). Some observations on psychosocial stressors, immunity, and individual differences in nonhuman primates. *American Journal of Primatology*, *39*(4), 205–221. https://doi.org/10.1002/(SICI)1098-2345(1996)39:4<205::AID-AJP2>3.0.CO;2-2

Lauriola, M., & Levin, I. P. (2001). Personality traits and risky decision-making in a controlled experimental task: An exploratory study. *Personality and Individual Differences*, *31*(2), 215–226. https://doi.org/10.1016/S0191-8869(00)00130-6

Leeuwen, E. J. C. van, Cronin, K. A., & Haun, D. B. M. (2018). Population-specific social dynamics in chimpanzees. *Proceedings of the National Academy of Sciences*, *115*(45), 11393–11400. https://doi.org/10.1073/pnas.1722614115

Lefevre, C. E., Wilson, V. A. D., Morton, F. B., Brosnan, S. F., Paukner, A., & Bates, T. C. (2014). Facial Width-To-Height Ratio Relates to Alpha Status and Assertive Personality in Capuchin Monkeys. *PLOS ONE*, *9*(4), e93369. https://doi.org/10.1371/journal.pone.0093369

Lehmann, J., & Boesch, C. (2009). Sociality of the dispersing sex: The nature of social bonds in West African female chimpanzees, Pan troglodytes. *Animal Behaviour*, *77*(2), 377–387. https://doi.org/10.1016/j.anbehav.2008.09.038

Lilienfeld, S. O., Gershon, J., Duke, M., Marino, L., & de Waal, F. B. M. (1999). A preliminary investigation of the construct of psychopathic personality (psychopathy) in chimpanzees (Pan troglodytes). *Journal of Comparative Psychology*, *113*(4), 365–375. https://doi.org/10.1037/0735-7036.113.4.365

Locke, D. P., Hillier, L. W., Warren, W. C., Worley, K. C., Nazareth, L. V., Muzny, D. M., … Wilson, R. K. (2011). Comparative and demographic analysis of orang-utan genomes. *Nature*, *469*(7331), 529–533. https://doi.org/10.1038/nature09687

Locke, K. D., Locke, E. A., Morgan, G. A., & Zimmermann, R. R. (1964). Dimensions of Social Interactions among Infant Rhesus Monkeys. *Psychological Reports*, *15*(2), 339–349. https://doi.org/10.2466/pr0.1964.15.2.339

Locke, K. D., Morgan, G. A., & Zimmermann, R. R. (1964). Method for Observing Social Interactions in Groups of Infant Rhesus Monkeys. *Psychological Reports*, *14*(1), 83–91. https://doi.org/10.2466/pr0.1964.14.1.83

Loewen, S., & Gonulal, T. (2015). Exploratory Factor Analysis and Principal Components Analysis. In L. Plonsky (Ed.), *Advancing Quantitative Methods in Second Language Research* (1st ed., pp. 182–212; By L. Plonsky). https://doi.org/10.4324/9781315870908-9

Maestripieri, D. (2000). Measuring temperament in rhesus macaques: Consistency and change in emotionality over time. *Behavioural Processes*, *49*(3), 167–171. https://doi.org/10.1016/S0376-6357(00)00083-8

Maestripieri, D., & Hoffman, C. L. (2011). Chronic stress, allostatic load, and aging in nonhuman primates. *Development and Psychopathology*, *23*(4), 1187–1195. https://doi.org/10.1017/S0954579411000551

Manson, J. H., & Perry, S. (2000). Correlates of Self-Directed Behaviour in Wild White-Faced Capuchins. *Ethology*, *106*(4), 301–317. https://doi.org/10.1046/j.1439-0310.2000.00527.x

Manson, J. H., & Perry, S. (2013). Personality structure, sex differences, and temporal change and stability in wild white-faced capuchins (Cebus capucinus). *Journal of Comparative Psychology*, *127*(3), 299–311. https://doi.org/10.1037/a0031316

Marrus, N., Faughn, C., Shuman, J., Petersen, S. E., Constantino, J. N., Povinelli, D. J., & Pruett, J. R. (2011). Initial Description of a Quantitative, Cross-Species (Chimpanzee–Human) Social Responsiveness Measure. *Journal of the American Academy of Child & Adolescent Psychiatry*, *50*(5), 508–518. https://doi.org/10.1016/j.jaac.2011.01.009

Martau, P. A., Caine, N. G., & Candland, D. K. (1985). Reliability of the emotions profile index, primate form, withPapio hamadryas, Macaca fuscata, and twoSaimiri species. *Primates*, *26*(4), 501–505. https://doi.org/10.1007/BF02382466

Martin, J. E. (2005). The influence of rearing on personality ratings of captive chimpanzees (Pan troglodytes). *Applied Animal Behaviour Science*, *90*(2), 167–181. https://doi.org/10.1016/j.applanim.2004.08.019

Martin, J. G. A., Nussey, D. H., Wilson, A. J., & Réale, D. (2011). Measuring individual differences in reaction norms in field and experimental studies: A power analysis of random regression models. *Methods in Ecology and Evolution*, *2*(4), 362–374. https://doi.org/10.1111/j.2041-210X.2010.00084.x

Martin, J. S., & Suarez, S. A. (2017). Personality assessment and model comparison with behavioral data: A statistical framework and empirical demonstration with bonobos (*Pan paniscus*). *American Journal of Primatology*, *79*(8), e22670. https://doi.org/10.1002/ajp.22670

Martin, P., & Bateson, P. P. G. (2007). *Measuring behaviour: An introductory guide* (3rd ed). Cambridge ; New York: Cambridge University Press.

Marzec, A. M., Kunz, J. A., Falkner, S., Atmoko, S. S. U., Alavi, S. E., Moldawer, A. M., … van Noordwijk, M. A. (2016). The dark side of the red ape: Male-mediated lethal female competition in Bornean

orangutans. *Behavioral Ecology and Sociobiology*, *70*(4), 459–466. https://doi.org/10.1007/s00265-015-2053-3

Masi, S., Gustafsson, E., Saint Jalme, M., Narat, V., Todd, A., Bomsel, M.-C., & Krief, S. (2012). Unusual feeding behavior in wild great apes, a window to understand origins of self-medication in humans: Role of sociality and physiology on learning process. *Physiology & Behavior*, *105*(2), 337–349. https://doi.org/10.1016/j.physbeh.2011.08.012

Massen, J. J. M., Antonides, A., Arnold, A.-M. K., Bionda, T., & Koski, S. E. (2013). A behavioral view on chimpanzee personality: Exploration tendency, persistence, boldness, and tool-orientation measured with group experiments: Behavioral Measures of Chimp Personality. *American Journal of Primatology*, *75*(9), 947–958. https://doi.org/10.1002/ajp.22159

Massen, J. J. M., & Koski, S. E. (2014). Chimps of a feather sit together: Chimpanzee friendships are based on homophily in personality. *Evolution and Human Behavior*, *35*(1), 1–8. https://doi.org/10.1016/j.evolhumbehav.2013.08.008

Matsuzawa, T., Biro, D., Humle, T., Inoue-Nakamura, N., Tonooka, R., & Yamakoshi, G. (2001). Emergence of Culture in Wild Chimpanzees: Education by Master-Apprenticeship. In T. Matsuzawa (Ed.), *Primate Origins of Human Cognition and Behavior* (pp. 557–574). https://doi.org/10.1007/978-4-431-09423-4_28

McCowan, B., Beisner, B. A., Capitanio, J. P., Jackson, M. E., Cameron, A. N., Seil, S., … Fushing, H. (2011). Network Stability Is a Balancing Act of Personality, Power, and Conflict Dynamics in Rhesus Macaque Societies. *PLOS ONE*, *6*(8), e22350. https://doi.org/10.1371/journal.pone.0022350

McCrae, R. R., & Costa, P. T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology*, *52*(1), 81–90. https://doi.org/10.1037//0022-3514.52.1.81

McCrae, R. R., & Costa, P. T. (1997). Personality trait structure as a human universal. *American Psychologist*, *52*(5), 509–516. https://doi.org/10.1037/0003-066X.52.5.509

McCrae, Robert R., & John, O. P. (1992). An Introduction to the Five-Factor Model and Its Applications. *Journal of Personality*, *60*(2), 175–215. https://doi.org/10.1111/j.1467-6494.1992.tb00970.x

McDonald, J. D. (2008). Measuring Personality Constructs: The Advantages and Disadvantages of Self-Reports, Informant Reports and Behavioural Assessments. *Enquire*, *1*(1), 1–19.

McDougall, P. T., Réale, D., Sol, D., & Reader, S. M. (2006). Wildlife conservation and animal temperament: Causes and consequences of evolutionary change for captive, reintroduced, and wild populations. *Animal Conservation*, *9*(1), 39–48. https://doi.org/10.1111/j.1469-1795.2005.00004.x

McFarland, R., Murphy, D., Lusseau, D., Henzi, S. P., Parker, J. L., Pollet, T. V., & Barrett, L. (2017). The 'strength of weak ties' among female baboons: Fitness-related benefits of social bonds. *Animal Behaviour*, *126*, 101–106. https://doi.org/10.1016/j.anbehav.2017.02.002

McGrew, W. C., McGrew, W. C., Marchant, L. F., & Nishida, T. (1996). *Great Ape Societies*. Cambridge University Press.

McGuire, M. T., Raleigh, M. J., & Pollack, D. B. (1994). Personality features in vervet monkeys: The effects of sex, age, social status, and group composition. *American Journal of Primatology*, *33*(1), 1–13. https://doi.org/10.1002/ajp.1350330102

McLennan, M. R., Spagnoletti, N., & Hockings, K. J. (2017). The Implications of Primate Behavioral Flexibility for Sustainable Human–Primate Coexistence in Anthropogenic Habitats. *International Journal of Primatology*, *38*(2), 105–121. https://doi.org/10.1007/s10764-017-9962-0

Mehl, M. R., Gosling, S. D., & Pennebaker, J. W. (2006). Personality in its natural habitat: Manifestations and implicit folk theories of personality in daily life. *Journal of Personality and Social Psychology*, *90*(5), 862–877. https://doi.org/10.1037/0022-3514.90.5.862

Mehta, P. H., & Gosling, S. D. (2008). Bridging human and animal research: A comparative approach to studies of personality and health. *Brain, Behavior, and Immunity*, *22*(5), 651–661. https://doi.org/10.1016/j.bbi.2008.01.008

Mendoza, S. P., Capitanio, J. P., & Mason, W. (2000). Chronic social stress: Studies in non-human primates. In G. P. Moberg & J. A. Mench, *The Biology of Animal Stress Basic Principles and Implications for Animal Welfare* (pp. 227–248). Wallingford, UK: CABI Publishing.

Mielke, A., Preis, A., Samuni, L., Gogarten, J. F., Wittig, R. M., & Crockford, C. (2018). Flexible decision-making in grooming partner choice in sooty mangabeys and chimpanzees. *Royal Society Open Science*, *5*(7), 172143. https://doi.org/10.1098/rsos.172143

Miller, D. B. (1977). Roles of naturalistic observation in comparative psychology. *American Psychologist*, *32*(3), 211–219. https://doi.org/10.1037/0003-066X.32.3.211

Mineka, S., Keir, R., & Price, V. (1980). Fear of snakes in wild- and laboratory-reared rhesus monkeys (Macaca mulatta). *Animal Learning & Behavior*, *8*(4), 653–663. https://doi.org/10.3758/BF03197783

Mischel, W. (1968). *Personality and assessment.* New York: Wiley.

Mischel, W., & Peake, P. K. (1982). Beyond Deja Vu in the Search for Cross-Situational Consistency. *Psychological Review*, *89*(6), 730–755.

Mischel, W., & Shoda, Y. (1995). A cognitive-affective system theory of personality: Reconceptualizing situations, dispositions, dynamics, and invariance in personality structure. *Psychological Review*, *102*(2), 246–268. https://doi.org/10.1037/0033-295X.102.2.246

Mischel, W., Shoda, Y., & Mendoza-Denton, R. (2002). Situation-Behavior Profiles as a Locus of Consistency in Personality. *Current Directions in Psychological Science*, *11*(2), 50.

Mitani, J. C. (2009). Male chimpanzees form enduring and equitable social bonds. *Animal Behaviour*, *77*(3), 633–640. https://doi.org/10.1016/j.anbehav.2008.11.021

Mitani, J. C., & Watts, D. P. (2001). Why do chimpanzees hunt and share meat? *Animal Behaviour*, *61*(5), 915–924. https://doi.org/10.1006/anbe.2000.1681

Mitani, J. C., Watts, D. P., & Muller, M. N. (2002). Recent developments in the study of wild chimpanzee behavior. *Evolutionary Anthropology: Issues, News, and Reviews*, *11*(1), 9–25. https://doi.org/10.1002/evan.10008

Mondragon-Ceballos, R., & Santillán-Doherty, A. M. (1994). The relationship between personality and age, sex and rank in captive stumptail macaques. In B. Roeder, B. Thierry, S. Anderson, & N. Herrenschmidt (Eds.), *Current primatology, vol. II: social development, learning and behavior.* (pp. 241–250). Strasbourg: University of Louis Pasteur.

Morton, F. B., Lee, P. C., Buchanan-Smith, H. M., Brosnan, S. F., Thierry, B., Paukner, A., … Weiss, A. (2013). Personality structure in brown capuchin monkeys (Sapajus apella): Comparisons with chimpanzees (Pan troglodytes), orangutans (Pongo spp.), and rhesus macaques (Macaca mulatta). *Journal of Comparative Psychology*, *127*(3), 282–298. https://doi.org/10.1037/a0031723

Moskowitz, D. S. (1982). Coherence and cross-situational generality in personality: A new analysis of old problems. *Journal of Personality and Social Psychology*, *43*(4), 754.

Mullen, P. E. (1996). Jealousy and the emergence of violent and intimidating behaviours Editorial. *Criminal Behaviour and Mental Health*, (3), 199–206.

Muller, M. N., & Mitani, J. C. (2005). Conflict and Cooperation in Wild Chimpanzees. In *Advances in the Study of Behavior* (Vol. 35, pp. 275–331). https://doi.org/10.1016/S0065-3454(05)35007-8

Murray, L. E. (1998). The effects of group structure and rearing strategy on personality in Chimpanzees Pan troglodytes at Chester, London ZSL and Twycross Zoos. *International Zoo Yearbook*, *36*(1), 97–108.

Mushquash, C., & O'Connor, B. P. (2006). SPSS and SAS programs for generalizability theory analyses. *Behavior Research Methods*, *38*(3), 542–547. https://doi.org/10.3758/BF03192810

Nettle, D. (2006). The evolution of personality variation in humans and other animals. *American Psychologist*, *61*(6), 622–631. https://doi.org/10.1037/0003-066X.61.6.622

Neumann, C., Agil, M., Widdig, A., & Engelhardt, A. (2013). Personality of wild male crested macaques (Macaca nigra). *PloS One*, *8*(8), e69383.

Nishida, T. (1990). *Chimpanzees of the Mahale Mountains*. Retrieved from http://agris.fao.org/agris-search/search.do?recordID=US201300184902

Nishida, T., Kano, T., Goodall, J., McGrew, W. C., & Nakamura, M. (1999). Ethogram and Ethnography of Mahale Chimpanzees. *Anthropological Science*, *116*(2), 141–18. https://doi.org/10.1537/ase.107.141

Palagi, E., Cordoni, G., & Borgognini Tarli, S. M. (2004). Immediate and Delayed Benefits of Play Behaviour: New Evidence from Chimpanzees (*Pan troglodytes*). *Ethology*, *110*(12), 949–962. https://doi.org/10.1111/j.1439-0310.2004.01035.x

Palagi, E., & Paoli, T. (2007). Play in adult bonobos (Pan paniscus): Modality and potential meaning. *American Journal of Physical Anthropology*, *134*(2), 219–225. https://doi.org/10.1002/ajpa.20657

Parish, A. R. (1996). Female relationships in bonobos(Pan paniscus). *Hu Nat*, *7*(1), 61–96. https://doi.org/10.1007/BF02733490

Paunonen, S. V. (2001). Inconsistencies in the Personality Consistency Debate. *Psychological Inquiry*, *12*(2), 91–93. Retrieved from JSTOR.

Pavani, S., Maestripieri, D., Schino, G., Turillazzi, P. G., & Scucchi, S. (1991). Factors influencing scratching behaviour in long-tailed macaques (Macaca fascicularis). *Folia Primatologica*, *57*(1), 34–38. https://doi.org/10.1159/000156561

Pavlov, I. (1906). The scientific investigation of the psychical faculties or processes in higher animals. *Science*, *24*, 613–619.

Pederson, A. K., King, J. E., & Landau, V. I. (2005). Chimpanzee (Pan troglodytes) personality predicts behavior. *Journal of Research in Personality*, *39*(5), 534–549. https://doi.org/10.1016/j.jrp.2004.07.002

Pellis, S. M., & Iwaniuk, A. N. (2000). Adult–Adult Play in Primates: Comparative Analyses of its Origin, Distribution and Evolution. *Ethology*, *106*(12), 1083–1104. https://doi.org/10.1046/j.1439-0310.2000.00627.x

Penn, D. J., & Frommen, J. G. (2010). Kin recognition: An overview of conceptual issues, mechanisms and evolutionary theory. In P. Kappeler (Ed.), *Animal Behaviour: Evolution and Mechanisms* (pp. 55–85). https://doi.org/10.1007/978-3-642-02624-9_3

Perelman, P., Johnson, W. E., Roos, C., Seuánez, H. N., Horvath, J. E., Moreira, M. A., ... & Schneider, M. P. C. (2011). A molecular phylogeny of living primates. *PLoS Genet*, 7(3), e1001342.

Perry, S. (1998). A case report of a male rank reversal in a group of wild white-faced capuchins (Cebus capucinus). *Primates*, *39*(1), 51–70. https://doi.org/10.1007/BF02557743

Pervin, L. A., & John, O. P. (1997). *Personality: Theory and research*. New York, US: John Wiley.

Peterson, R. A. (1994). A Meta-analysis of Cronbach's Coefficient Alpha. *Journal of Consumer Research*, *21*(2), 381–391. https://doi.org/10.1086/209405

Poirier, F. E., Bellisari, A., & Haines, L. (1978). Functions of primate play behavior. In E. O. Smith (Ed.), *Social play in primates* (pp. 143–168). New York, US: Academic Press.

Preuschoft, S., & van Hooff, J. A. (1995). Homologizing primate facial displays: A critical review of methods. *Folia Primatologica; International Journal of Primatology*, *65*(3), 121–137.

Pritchard, A., Sheeran, L. K., Gabriel, K. I., Li, J.-H., & Wagner, R. S. (2014). Behaviors that predict personality components in adult free-ranging Tibetan macaques Macaca thibetana. *Current Zoology*, *60*(3), 362–372.

Putman, R. J. (1995). Ethical considerations and animal welfare in ecological field studies. *Biodiversity & Conservation*, *4*(8), 903–915. https://doi.org/10.1007/BF00056197

Quenette, P. Y. (1990). Functions of vigilance behaviour in mammals: A review. *Acta Oecologica*, *11*(6), 801–818.

Ramakrishnan, U., Coss, R. G., Schank, J., Dharawat, A., & Kim, S. (2005). Snake Species Discrimination by Wild Bonnet Macaques (Macaca radiata). *Ethology*, *111*(4), 337–356. https://doi.org/10.1111/j.1439-0310.2004.01063.x

Réale, D., Gallant, B. Y., Leblanc, M., & Festa-Bianchet, M. (2000a). Consistency of temperament in bighorn ewes and correlates with behaviour and life history. *Animal Behaviour*, *60*(5), 589–597. https://doi.org/10.1006/anbe.2000.1530

Réale, D., Gallant, B. Y., Leblanc, M., & Festa-Bianchet, M. (2000b). Consistency of temperament in bighorn ewes and correlates with behaviour and life history. *Animal Behaviour*, *60*(5), 589–597. https://doi.org/10.1006/anbe.2000.1530

Réale, D., Reader, S. M., Sol, D., McDougall, P. T., & Dingemanse, N. J. (2007). Integrating animal temperament within ecology and evolution. *Biological Reviews*, *82*(2), 291–318. https://doi.org/10.1111/j.1469-185X.2007.00010.x

Reaney, L. T., & Backwell, P. R. Y. (2007). Risk-taking behavior predicts aggression and mating success in a fiddler crab. *Behavioral Ecology*, *18*(3), 521–525. https://doi.org/10.1093/beheco/arm014

Reite, M., & Short, R. (1980). A biobehavioral developmental profile (BDP) for the pigtailed monkey. *Developmental Psychobiology*, *13*(3), 243–285. https://doi.org/10.1002/dev.420130303

Revelle, W. (2019). *Package 'psych'*. Retrieved from https://cran.r-project.org/web/packages/psych/psych.pdf

Roberts, B. W., & DelVecchio, W. F. (2000). The rank-order consistency of personality traits from childhood to old age: a quantitative review of longitudinal studies. *Psychological bulletin*, 126(1), 3.

Robbins, M. M., Ando, C., Fawcett, K. A., Grueter, C. C., Hedwig, D., Iwata, Y., … Yamagiwa, J. (2016). Behavioral Variation in Gorillas: Evidence of Potential Cultural Traits. *PLOS ONE*, *11*(9), e0160483. https://doi.org/10.1371/journal.pone.0160483

Robinson, L. M., Altschul, D. M., Wallace, E. K., Úbeda, Y., Llorente, M., Machanda, Z., … Weiss, A. (2017). Chimpanzees with positive welfare are happier, extraverted, and emotionally stable. *Applied Animal Behaviour Science*, *191*, 90–97. https://doi.org/10.1016/j.applanim.2017.02.008

Robinson, L. M., Coleman, K., Capitanio, J. P., Gottlieb, D. H., Handel, I. G., Adams, M. J., … Weiss, A. (2018). Rhesus macaque personality, dominance, behavior, and health. *American Journal of Primatology*, *80*(2), e22739. https://doi.org/10.1002/ajp.22739

Rogers, J., Shelton, S. E., Shelledy, W., Garcia, R., & Kalin, N. H. (2008). Genetic influences on behavioral inhibition and anxiety in juvenile rhesus macaques. *Genes, Brain and Behavior*, *7*(4), 463–469. https://doi.org/10.1111/j.1601-183X.2007.00381.x

Rommeck, I., Capitanio, J. P., Strand, S. C., & McCowan, B. (2011). Early social experience affects behavioral and physiological responsiveness to stressful conditions in infant rhesus macaques

(Macaca mulatta). *American Journal of Primatology*, *73*(7), 692–701. https://doi.org/10.1002/ajp.20953

Rouff, J. H., Sussman, R. W., & Strube, M. J. (2005). Personality traits in captive lion-tailed macaques (Macaca silenus). *American Journal of Primatology*, *67*(2), 177–198. https://doi.org/10.1002/ajp.20176

Russon, A. (2006). Acquisition of Complex Foraging Skills in Juvenile and Adolescent Orangutans (Pongo pygmaeus): Developmental Influences. *Aquatic Mammals*, *32*(4), 500–510. https://doi.org/10.1578/AM.32.4.2006.500

Russon, A., Kuncoro, P., Ferisa, A., & Putri Handayani, D. (2010). How Orangutans (Pongo pygmaeus) Innovate for Water. *Journal of Comparative Psychology (Washington, D.C. : 1983)*, *124*, 14–28. https://doi.org/10.1037/a0017929

Santillán-Doherty, A. M., Cortés-Sotres, J., Arenas-Rosas, R. V., Márquez-Arias, A., Cruz, C., Medellín, A., … Díaz, J. L. (2010). Novelty-Seeking Temperament in Captive Stumptail Macaques (Maca ea arctoides) and Spider Monkeys (Ateles geoffroyi). *Loumal al Comparative Psychology*, *124*(2), 211–218. https://doi.org/10.1O37/aOOI8267

Santillán-Doherty, A. M., Muñóz-Delgado, J., Arenas, R., Márquez, A., & Cortés, J. (2006). Reliability of a method to measure novelty-seeking in nonhuman primates. *American Journal of Primatology*, *68*(11), 1098–1113. https://doi.org/10.1002/ajp.20310

Sapolsky, R. M., & Ray, J. C. (1989). Styles of dominance and their endocrine correlates among wild olive baboons (Papio anubis). *American Journal of Primatology*, *18*(1), 1–13. https://doi.org/10.1002/ajp.1350180102

Sapolsky, R. M., & Share, L. J. (2004). A Pacific Culture among Wild Baboons: Its Emergence and Transmission. *PLOS Biology*, *2*(4), e106. https://doi.org/10.1371/journal.pbio.0020106

Schaefer, S. A., & Steklis, H. D. (2014). Personality and subjective well-being in captive male western lowland gorillas living in bachelor groups: Personality and Well-Being in Gorillas. *American Journal of Primatology*, *76*(9), 879–889. https://doi.org/10.1002/ajp.22275

Schneider, M. L., Moore, C. F., Suomi, S. J., & Champoux, M. (1991). Laboratory Assessment of Temperament and Environmental Enrichment in Rhesus Monkey Infants (Macaca mulatta). *American Journal of Primatology*, *25*(3), 137–155. https://doi.org/10.1002/ajp.1350250302

Schuppli, C., Forss, S. I. F., Meulman, E. J. M., Zweifel, N., Lee, K. C., Rukmana, E., … van Schaik, C. P. (2016). Development of foraging skills in two orangutan populations: Needing to learn or needing to grow? *Frontiers in Zoology*, *13*(1), 43. https://doi.org/10.1186/s12983-016-0178-5

Schuppli, C., Forss, S., Meulman, E., Atmoko, S. U., Noordwijk, M. van, & Schaik, C. van. (2017). The effects of sociability on exploratory tendency and innovation repertoires in wild Sumatran and Bornean orangutans. *Scientific Reports*, *7*(1), 15464. https://doi.org/10.1038/s41598-017-15640-x

Schuppli, C., Meulman, E. J. M., Forss, S. I. F., Aprilinayati, F., van Noordwijk, M. A., & van Schaik, C. P. (2016). Observational social learning and socially induced practice of routine skills in immature wild orang-utans. *Animal Behaviour*, *119*, 87–98. https://doi.org/10.1016/j.anbehav.2016.06.014

Schuster, A. C., Carl, T., & Foerster, K. (2017). Repeatability and consistency of individual behaviour in juvenile and adult Eurasian harvest mice. *Die Naturwissenschaften*, *104*(3). https://doi.org/10.1007/s00114-017-1430-3

Seyfarth, R. M., Silk, J. B., & Cheney, D. L. (2012). Variation in personality and fitness in wild female baboons. *Proceedings of the National Academy of Sciences*, *109*(42), 16980–16985. https://doi.org/10.1073/pnas.1210780109

Shavelson, R. J., & Webb, N. M. (1991). *Generalizability Theory: A Primer*. SAGE.

Shavelson, R. J., Webb, N. M., & Rowley, G. L. (1989). Generalizability theory. *American Psychologist*, *44*(6), 922–932. https://doi.org/10.1037/0003-066X.44.6.922

Shoda, Y., Mischel, W., & Wright, J. C. (1994). Intraindividual Stability in the Organization and Patterning of Behavior: Incorporating Psychological Situations into the Idiographic Analysis of Personality. *Journal of Personality & Social Psychology*, *67*(4), 674–687. https://doi.org/10.1037/0022-3514.67.4.674

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, *86*(2), 420–428. https://doi.org/10.1037/0033-2909.86.2.420

Shultz, S., Faurie, C., & Noë, R. (2003). Behavioural responses of Diana monkeys to male long-distance calls: Changes in ranging, association patterns and activity. *Behavioral Ecology and Sociobiology*, *53*(4), 238–245. https://doi.org/10.1007/s00265-002-0575-y

Shumaker, R. W., Wich, S. A., & Perkins, L. (2008). Reproductive Life History Traits of Female Orangutans (Pongo spp.). *Interdisciplinary Topics in Gerontology*, *36*, 147–161.

Sih, A., Bell, A., & Johnson, J. C. (2004). Behavioral syndromes: An ecological and evolutionary overview. *Trends in Ecology & Evolution*, *19*(7), 372–378. https://doi.org/10.1016/j.tree.2004.04.009

Sih, A., & Bell, A. M. (2008). Insights for Behavioral Ecology from Behavioral Syndromes. In T. J. R. H. Jane Brockmann (Ed.), *Advances in the Study of Behavior*: *Vol. Volume 38* (pp. 227–281). Retrieved from http://www.sciencedirect.com/science/article/pii/S0065345408000053

Sih, A., Bell, A. M., Johnson, J. C., & Ziemba, R. E. (2004). Behavioral Syndromes: An Integrative Overview. *The Quarterly Review of Biology*, *79*(3), 241–277. https://doi.org/10.1086/422893

Silk, J., Alberts, S. C., & Altmann, J. (2003). Social Bonds of Female Baboons Enhance Infant Survival. *Science*, *302*(5648), 1231–1234. https://doi.org/10.1126/science.1088580

Silk, J., Beehner, J. C., Bergman, T. J., Crockford, C., Engh, A. L., Moscovice, L. R., … Cheney, D. L. (2009). The benefits of social capital: Close social bonds among female baboons enhance offspring survival. *Proceedings of the Royal Society B: Biological Sciences*, *276*(1670), 3099–3104. https://doi.org/10.1098/rspb.2009.0681

Silk, J. B., Brosnan, S. F., Henrich, J., Lambeth, S. P., & Shapiro, S. (2013). Chimpanzees share food for many reasons: the role of kinship, reciprocity, social bonds and harassment on food transfers. *Animal behaviour*, 85(5), 941-947

Singleton, I., Knott, C. D., Morrogh-Bernard, H. C., Wich, S. A., van Schaik, C. P., Wich, S. A., … van Schaik, C. P. (2009). Ranging behavior of orangutan females and social organization. In *Orangutans: Geographic variation in behavioral ecology and conservation* (pp. 205–213). https://doi.org/10.1093/acprof:oso/9780199213276.003.0013

Šlipogor, V., Gunhold-de Oliveira, T., Tadić, Z., Massen, J. J. M., & Bugnyar, T. (2016). Consistent inter-individual differences in common marmosets (Callithrix jacchus) in Boldness-Shyness, Stress-Activity, and Exploration-Avoidance. *American Journal of Primatology*, *78*(9), 961–973. https://doi.org/10.1002/ajp.22566

Slocombe, K. E., & Zuberbühler, K. (2005). Agonistic Screams in Wild Chimpanzees (Pan troglodytes schweinfurthii) Vary as a Function of Social Role. *Journal of Comparative Psychology*, *119*(1), 67–77. https://doi.org/10.1037/0735-7036.119.1.67

Smith, B. R., & Blumstein, D. T. (2008). Fitness consequences of personality: A meta-analysis. *Behavioral Ecology*, *19*(2), 448–455. https://doi.org/10.1093/beheco/arm144

Smuts, B. B., Cheney, D. L., Seyfarth, R. M., & Wrangham, R. W. (2008). *Primate Societies*. University of Chicago Press.

Spencer-Booth, Y., & Hinde, R. A. (1969). Tests of Behavioural Characteristics for Rhesus Monkeys. *Behaviour*, *33*(3–4), 179–210. https://doi.org/10.1163/156853969X00062

Spinka, M., Newberry, R. C., & Bekoff, M. (2001). Mammalian Play: Training for the Unexpected. *The Quarterly Review of Biology*, *76*(2), 141–168. https://doi.org/10.1086/393866

Staes, N., Eens, M., Weiss, A., & Stevens, J. M. G. (2016). Bonobo personality: Age and sex effects and links with behavior and dominance. In B. Hare & S. Yamamoto, *Bonobos: Unique in Mind, Brain, and Behavior*. Retrieved from https://www.oxfordscholarship.com/view/10.1093/oso/9780198728511.001.0001/oso-9780198728511-chapter-13

Stauffer, R. L., Walker, A., Ryder, O. A., Lyons-Weiler, M., & Hedges, S. B. (2001). Human and Ape Molecular Clocks and Constraints on Paleontological Hypotheses. *Journal of Heredity*, *92*(6), 469–474. https://doi.org/10.1093/jhered/92.6.469

Stevenson-Hinde, J., Stillwell-Barnes, R., & Zunz, M. (1980a). Individual differences in young rhesus monkeys: Consistency and change—Springer. *Primates*, *21*(4), 498–509.

Stevenson-Hinde, J., Stillwell-Barnes, R., & Zunz, M. (1980b). Subjective assessment of rhesus monkeys over four successive years. *Primates*, *21*(1), 66–82. https://doi.org/10.1007/BF02383825

Stevenson-Hinde, J., & Zunz, M. (1978). Subjective assessment of individual rhesus monkeys. *Primates*, *19*(3), 473–482. https://doi.org/10.1007/BF02373309

Stevenson-Hinde, J., Zunz, M., & Stillwell-Barnes, R. (1980). Behaviour of one-year-old rhesus monkeys in a strange situation. *Animal Behaviour*, *28*(1), 266–277. https://doi.org/10.1016/S0003-3472(80)80029-7

Stoinski, T. S., Kuhar, C. W., Lukas, K. E., & Maple, T. L. (2004). Social dynamics of captive western lowland gorillas living in all-male groups. *Behaviour*, *141*(2), 169–195. https://doi.org/10.1163/156853904322890807

Suarez-Jimenez, B., Hathaway, A., Waters, C., Vaughan, K., Suomi, S. J., Noble, P. L., … Nelson, E. E. (2013). Effect of Mother's Dominance Rank on Offspring Temperament in Infant Rhesus Monkeys (Macaca mulatta). *American Journal of Primatology*, *75*(1), 65–73. https://doi.org/10.1002/ajp.22081

Sullivan, E. C., Mendoza, S. P., & Capitanio, J. P. (2011). Similarity in temperament between mother and offspring rhesus monkeys: Sex differences and the role of monoamine oxidase-a and serotonin transporter promoter polymorphism genotypes. *Developmental Psychobiology*, *53*(6), 549–563. https://doi.org/10.1002/dev.20594

Suomi, S. J., Novak, M. A., & Well, A. (1996). Aging in rhesus monkeys: Different windows on behavioral continuity and change. *Developmental Psychology*, *32*(6), 1116–1128. https://doi.org/10.1037/0012-1649.32.6.1116

Sussman, A. F., Ha, J. C., Bentson, K. L., & Crockett, C. M. (2013). Temperament in Rhesus, Long-Tailed, and Pigtailed Macaques Varies by Species and Sex. *American Journal of Primatology*, *75*, 303–313. https://doi.org/10.1002/ajp.22104

Sussman, A. F., Mates, E. A., Ha, J. C., Bentson, K. L., & Crockett, C. M. (2014). Tenure in current captive setting and age predict personality changes in adult pigtailed macaques. *Animal Behaviour*, *89*, 23–30. https://doi.org/10.1016/j.anbehav.2013.12.009

Sussman, A., & Ha, J. (2011). Developmental and cross-situational stability in infant pigtailed macaque temperament. *Developmental Psychology*, *47*(3), 781–791. https://doi.org/10.1037/a0022999

Tajfel, H. (1974). Social identity and intergroup behaviour. *Information (International Social Science Council)*, *13*(2), 65–93. https://doi.org/10.1177/053901847401300204

Tkaczynski, P. J., Ross, C., MacLarnon, A., Mouna, M., Majolo, B., & Lehmann, J. (2018). Measuring personality in the field: An In Situ comparison of personality quantification methods in wild Barbary macaques (Macaca sylvanus). *Journal of Comparative Psychology (Washington, D.C.: 1983)*. https://doi.org/10.1037/com0000163

Tomassetti, D., Caracciolo, S., Manciocco, A., Chiarotti, F., Vitale, A., & De Filippis, B. (2019). Personality and lateralization in common marmosets (Callithrix jacchus). *Behavioural Processes*, *167*, 103899. https://doi.org/10.1016/j.beproc.2019.103899

Triandis, H. C. (2001). Individualism-Collectivism and Personality. *Journal of Personality*, *69*(6), 907–924. https://doi.org/10.1111/1467-6494.696169

Triandis, H. C., & Suh, E. M. (2002). Cultural Influences on Personality. *Annual Review of Psychology*, *53*(1), 133–160. https://doi.org/10.1146/annurev.psych.53.100901.135200

Troisi, A., Schino, G., D'Antoni, M., Pandolfi, N., Aureli, F., & D'Amato, F. R. (1991). Scratching as a behavioral index of anxiety in macaque mothers. *Behavioral and Neural Biology*, *56*(3), 307–313. https://doi.org/10.1016/0163-1047(91)90469-7

Uher, J. (2008a). Comparative personality research: Methodological approaches. *European Journal of Personality*, *22*(5), 427–455. https://doi.org/10.1002/per.680

Uher, J. (2008b). Three methodological core issues of comparative personality research. *European Journal of Personality*, *22*(5), 475–496. https://doi.org/10.1002/per.688

Uher, J. (2011a). Individual behavioral phenotypes: An integrative meta-theoretical framework. Why "behavioral syndromes" are not analogs of "personality". *Developmental Psychobiology*, *53*(6), 521–548. https://doi.org/10.1002/dev.20544

Uher, J. (2011b). Personality in nonhuman primates: What can we learn from human personality psychology? In A. Weiss, J. King, & L. Murray (Eds.), *Personality and Temperament in Nonhuman Primates* (pp. 41–76). Retrieved from http://janauher.com/paper/uher_2011b_chapter.pdf

Uher, J., Addessi, E., & Visalberghi, E. (2013). Contextualised behavioural measurements of personality differences obtained in behavioural tests and social observations in adult capuchin monkeys (Cebus apella). *Journal of Research in Personality*, *47*(4), 427–444. https://doi.org/10.1016/j.jrp.2013.01.013

Uher, J., & Asendorpf, J. B. (2008). Personality assessment in the Great Apes: Comparing ecologically valid behavior measures, behavior ratings, and adjective ratings. *Journal of Research in Personality*, *42*, 821–838.

Uher, J., Asendorpf, J. B., & Call, J. (2008). Personality in the behaviour of great apes: Temporal stability, cross-situational consistency and coherence in response. *Animal Behaviour*, *75*(1), 99–112. https://doi.org/10.1016/j.anbehav.2007.04.018

Uher, J., & Visalberghi, E. (2016). Observations versus assessments of personality: A five-method multi-species study reveals numerous biases in ratings and methodological limitations of standardised assessments. *Journal of Research in Personality*, *61*, 61–79. https://doi.org/10.1016/j.jrp.2016.02.003

Uher, J., Werner, C. S., & Gosselt, K. (2013). From observations of individual behaviour to social representations of personality: Developmental pathways, attribution biases, and limitations of questionnaire methods. *Journal of Research in Personality*, *47*(5), 647–667. https://doi.org/10.1016/j.jrp.2013.03.006

Utami Atmoko, S. S., Singleton, I., van Noordwijk, M. A., van Schaik, C. P., Mitra Setia, T., Wich, S. A., … van Schaik, C. P. (2009). Male-male relationships in orangutans. In *Orangutans: Geographic variation in behavioral ecology and conservation* (pp. 225–233). https://doi.org/10.1093/acprof:oso/9780199213276.003.0015

van Adrichem, G. G. J., Utami, S. S., Wich, S. A., van Hooff, J. A. R. A. M., & Sterck, E. H. M. (2006). The development of wild immature Sumatran orangutans (Pongo abelii) at Ketambe. *Primates*, *47*(4), 300–309. https://doi.org/10.1007/s10329-006-0193-9

van Hooff, J. (1973). A structural analysis of the social behavior of a semi-captive group of chimpanzees. In *Social communication and movement* (von Cranach M, Vine I, p. pp 75–162). Retrieved from https://www.scienceopen.com/document?vid=d6ab71f8-f37b-4c39-adae-715a7ed7d3bd

van Noordwijk, M. A., Arora, N., Willems, E. P., Dunkel, L. P., Amda, R. N., Mardianah, N., … van Schaik, C. P. (2012). Female philopatry and its social benefits among Bornean orangutans. *Behavioral Ecology and Sociobiology*, *66*(6), 823–834. https://doi.org/10.1007/s00265-012-1330-7

van Noordwijk, M. A., & van Schaik, C. P. (2005). Development of ecological competence in Sumatran orangutans. *American Journal of Physical Anthropology*, *127*(1), 79–94. https://doi.org/10.1002/ajpa.10426

van Oers, K., de Jong, G., van Noordwijk, A. J., Kempenaers, B., & Drent, P. J. (2005). Contribution of Genetics to the Study of Animal Personalities: A Review of Case Studies. *Behaviour*, *142*(9/10), 1185–1206.

van Oers, K., & Sinn, D. L. (2013). Quantitative and Molecular Genetics of Animal Personality. In C. Carere & D. Maestripieri (Eds.), *Animal Personalities: Behavior, Physiology, and Evolution* (pp. 149–200). Chicago, London: University of Chicago Press.

van Schaik, C. P. (1999). The socioecology of fission-fusion sociality in Orangutans. *Primates*, *40*(1), 69–86. https://doi.org/10.1007/BF02557703

van Schaik, C. P. (2013). The costs and benefits of flexibility as an expression of behavioural plasticity: A primate perspective. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *368*(1618), 20120339. https://doi.org/10.1098/rstb.2012.0339

van Schaik, C. P., Burkart, J., Damerius, L., Forss, S. I. F., Koops, K., Noordwijk, M. A. van, & Schuppli, C. (2016). The reluctant innovator: Orangutans and the phylogeny of creativity. *Phil. Trans. R. Soc. B*, *371*(1690), 20150183. https://doi.org/10.1098/rstb.2015.0183

van Vianen, A. E. M., & De Dreu, C. K. W. (2001). Personality in teams: Its relationship to social cohesion, task cohesion, and team performance. *European Journal of Work and Organizational Psychology*, *10*(2), 97–120. https://doi.org/10.1080/13594320143000573

Vazire, S., Gosling, S. D., Dickey, A. S., & Schapiro, S. J. (2007). Measuring Personality in Nonhuman Animals. In R. W. Robins, R. C. Fraley, & R. F. Krueger (Eds.), *Handbook of Research Methods in Personality Psychology* (pp. 190–206). Guilford Press.

Vitale, A. F., Visalberghi, E., & De Lillo, C. (1991). Responses to a snake model in captive crab-eating macaques (Macaca fascicularis) and captive tufted capuchins (Cebus apella). *International Journal of Primatology*, *12*(3), 277–286. https://doi.org/10.1007/BF02547588

von Borell, C., Kulik, L., & Widdig, A. (2016). Growing into the self: The development of personality in rhesus macaques. *Animal Behaviour*, *122*, 183–195. https://doi.org/10.1016/j.anbehav.2016.10.013

Vonk, J., Weiss, A., & Kuczaj, S. A. (Eds.). (2017). Personality in nonhuman animals. Berlin, Germany: Springer International Publishing.

Waitt, C., & Buchanan-Smith, H. M. (2001). What time is feeding?: How delays and anticipation of feeding schedules affect stump-tailed macaque behavior. *Applied Animal Behaviour Science*, *75*(1), 75–85. https://doi.org/10.1016/S0168-1591(01)00174-5

Walumbwa, F. O., & Schaubroeck, J. (2009). Leader personality traits and employee voice behavior: Mediating roles of ethical leadership and work group psychological safety. *Journal of Applied Psychology*, *94*(5), 1275–1286. https://doi.org/10.1037/a0015848

Waterson, R. H., Lander, E. S., Wilson, R. K., & The Chimpanzee Sequencing and Analysis Consortium. (2005). Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature*, *437*(7055), 69–87. https://doi.org/10.1038/nature04072

Watson, S. K., Townsend, S. W., Schel, A. M., Wilke, C., Wallace, E. K., Cheng, L., … Slocombe, K. E. (2015). Vocal Learning in the Functionally Referential Food Grunts of Chimpanzees. *Current Biology*, *25*(4), 495–499. https://doi.org/10.1016/j.cub.2014.12.032

Watson, S., & Ward, J. (1996). Temperament and problem solving in the small-eared bushbaby (Otolemur garnettii). *JOURNAL OF COMPARATIVE PSYCHOLOGY*, *110*(4), 377–385.

Watts, D. P. (1998). Coalitionary mate guarding by male chimpanzees at Ngogo, Kibale National Park, Uganda. *Behavioral Ecology and Sociobiology*, *44*(1), 43–55. https://doi.org/10.1007/s002650050513

Webster, M. M., & Ward, A. J. W. (2011). Personality and social context. *Biological Reviews*, *86*(4), 759–773. https://doi.org/10.1111/j.1469-185X.2010.00169.x

Weinstein, T. A. R., & Capitanio, J. P. (2012). Longitudinal Stability of Friendships in Rhesus Monkeys (Macaca mulatta): Individual- and Relationship-level Effects. *Journal of Comparative Psychology (Washington, D.C. : 1983)*, *126*(1), 97–108. https://doi.org/10.1037/a0025607

Weinstein, T. A. R., Capitanio, J. P., & Gosling, S. D. (2008). Personality in Animals. In *Handbook of Personality, Third Edition: Theory and Research*. Guilford Press.

Weisbuch, M., Slepian, M. L., Clarke, A., Ambady, N., & Veenstra-VanderWeele, J. (2010). Behavioral Stability Across Time and Situations: Nonverbal Versus Verbal Consistency. *Journal of Nonverbal Behavior*, *34*(1), 43–56. https://doi.org/10.1007/s10919-009-0079-9

Weiss, A. (2017). Exploring Factor Space (and Other Adventures) with the Hominoid Personality Questionnaire. In J. Vonk, A. Weiss, & S. A. Kuczaj (Eds.), *Personality in Nonhuman Animals* (pp. 19–38). https://doi.org/10.1007/978-3-319-59300-5_2

Weiss, A., Adams, M. J., & Johnson, W. (2011). The big none: No evidence for a general factor of personality in chimpanzees, orangutans, or rhesus macaques. *Journal of Research in Personality*, *45*(4), 393–397. https://doi.org/10.1016/j.jrp.2011.04.006

Weiss, A., Adams, M. J., Widdig, A., & Gerald, M. S. (2011). Rhesus Macaques (Macaca mulatta) as Living Fossils of Hominoid Personality and Subjective Well-being. *Journal of Comparative Psychology (Washington, D.C. : 1983)*, *125*(1), 72–83. https://doi.org/10.1037/a0021187

Weiss, A., Adams Mark J., & King James E. (2011). Happy orang-utans live longer lives. *Biology Letters*, *7*(6), 872–874. https://doi.org/10.1098/rsbl.2011.0543

Weiss, A., Gartner, M. C., Gold, K. C., & Stoinski, T. S. (2013). Extraversion predicts longer survival in gorillas: An 18-year longitudinal study. *Proceedings of the Royal Society B: Biological Sciences*, *280*(1752), 20122231. https://doi.org/10.1098/rspb.2012.2231

Weiss, A., Inoue-Murayama, M., Hong, K.-W., Inoue, E., Udono, T., Ochiai, T., … King, J. E. (2009). Assessing chimpanzee personality and subjective well-being in Japan. *American Journal of Primatology*, *71*(4), 283–292. https://doi.org/10.1002/ajp.20649

Weiss, A., Inoue-Murayama, M., King, J. E., Adams, M. J., & Matsuzawa, T. (2012). All too human? Chimpanzee and orang-utan personalities are not anthropomorphic projections. *Animal Behaviour*, *83*(6), 1355–1365. https://doi.org/10.1016/j.anbehav.2012.02.024

Weiss, A., & King, J. E. (2015). Great ape origins of personality maturation and sex differences: A study of orangutans and chimpanzees. *Journal of Personality and Social Psychology*, *108*(4), 648–664. https://doi.org/10.1037/pspp0000022

Weiss, A., King, J. E., & Enns, R. M. (2002). Subjective well-being is heritable and genetically correlated with dominance in chimpanzees (Pan troglodytes). *Journal of Personality and Social Psychology*, *83*(5), 1141–1149. https://doi.org/10.1037/0022-3514.83.5.1141

Weiss, A., King, J. E., & Figueredo, A. J. (2000). The Heritability of Personality Factors in Chimpanzees (Pan troglodytes). *Behavior Genetics*, *30*(3), 213–221. https://doi.org/10.1023/A:1001966224914

Weiss, A., King, J. E., & Hopkins, W. D. (2007). A cross-setting study of chimpanzee (Pan troglodytes) personality structure and development: Zoological parks and Yerkes National Primate Research Center. *American Journal of Primatology*, *69*(11), 1264–1277. https://doi.org/10.1002/ajp.20428

Weiss, A., King, J. E., & Murray, L. (Eds.). (2011). Personality and temperament in nonhuman primates. Springer Science & Business Media.

Weiss, A., King, J. E., & Perkins, L. (2006). Personality and subjective well-being in orangutans (Pongo pygmaeus and Pongo abelii). *Journal of Personality and Social Psychology*, *90*(3), 501–511. https://doi.org/10.1037/0022-3514.90.3.501

Weiss, A., Staes, N., Pereboom, J. J. M., Inoue-Murayama, M., Stevens, J. M. G., & Eens, M. (2015). Personality in Bonobos. *Psychological Science*, *26*(9), 1430–1439. https://doi.org/10.1177/0956797615589933

Weiss, A., Wilson, M. L., Collins, D. A., Mjungu, D., Kamenya, S., Foerster, S., & Pusey, A. E. (2017). Personality in the chimpanzees of Gombe National Park. *Scientific Data*, *4*, 170146. https://doi.org/10.1038/sdata.2017.146

Wergård, E.-M., Westlund, K., Spångberg, M., Fredlund, H., & Forkman, B. (2016). Training success in group-housed long-tailed macaques (Macaca fascicularis) is better explained by personality than by social rank. *Applied Animal Behaviour Science*, *177*, 52–58. https://doi.org/10.1016/j.applanim.2016.01.017

Westergaard, G. C., Mehlman, P. T., Westergaard, G. C., Suomi, S. J., & Higley, J. D. (1999). CSF 5-HIAA and aggression in female macaque monkeys: Species and interindividual differences. *Psychopharmacology*, *146*(4), 440–446. https://doi.org/10.1007/PL00005489

Whiten, A. (2000). Primate Culture and Social Learning. *Cognitive Science*, *24*(3), 477–508. https://doi.org/10.1207/s15516709cog2403_6

Wich, S. A., Atmoko, uci S. U., Setia, T. M., & Schaik, C. P. van. (2010). *Orangutans: Geographic Variation in Behavioral Ecology and Conservation*. OUP Oxford.

Wich, S. A., Utami-Atmoko, S. S., Setia, T. M., Rijksen, H. D., Schürmann, C., van Hooff, J. A. R. A. M., & van Schaik, C. P. (2004). Life history of wild Sumatran orangutans (Pongo abelii). *Journal of Human Evolution*, *47*(6), 385–398. https://doi.org/10.1016/j.jhevol.2004.08.006

Williamson, D. E., Coleman, K., Bacanu, S.-A., Devlin, B. J., Rogers, J., Ryan, N. D., & Cameron, J. L. (2003). Heritability of fearful-anxious endophenotypes in infant rhesus macaques: A preliminary report. *Biological Psychiatry*, *53*(4), 284–291. https://doi.org/10.1016/S0006-3223(02)01601-3

Wilson, D. S., Clark, A. B., Coleman, K., & Dearstyne, T. (1994). Shyness and boldness in humans and other animals. *Trends in Ecology & Evolution*, *9*(11), 442–446. https://doi.org/10.1016/0169-5347(94)90134-1

Winter, D. G., & Barenbaum, N. B. (1999). History of modern personality theory and research. In *Handbook of personality: Theory and research, 2nd ed* (pp. 3–27). New York, NY, US: Guilford Press.

Wolf, M., & Weissing, F. J. (2010). An explanatory framework for adaptive personality differences. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1560), 3959–3968. https://doi.org/10.1098/rstb.2010.0215

Wolf, M., & Weissing, F. J. (2012). Animal personalities: Consequences for ecology and evolution. *Trends in Ecology & Evolution*, *27*(8), 452–461. https://doi.org/10.1016/j.tree.2012.05.001

Wu, K. D., & Clark, L. A. (2003). Relations between personality traits and self-reports of daily behavior. *Journal of Research in Personality*, *37*(4), 231–256. https://doi.org/10.1016/S0092-6566(02)00539-1

Yamanashi, Y., Nogami, E., Teramoto, M., Morimura, N., & Hirata, S. (2018). Adult-adult social play in captive chimpanzees: Is it indicative of positive animal welfare? *Applied Animal Behaviour Science*, *199*, 75–83. https://doi.org/10.1016/j.applanim.2017.10.006

Yerkes, R. M. (1940). Social behavior of chimpanzees: Dominance between mates, in relation to sexual status. *Journal of Comparative Psychology*, *30*(1), 147–186. https://doi.org/10.1037/h0059232

Yerkes, R. M., & Yerkes, A. W. (1936). Nature and conditions of avoidance (fear) response in chimpanzee. *Journal of Comparative Psychology*, *21*(1), 53–66. https://doi.org/10.1037/h0058825

Appendices

A. Appendix 1. Overview of research on nonhuman primate personality

Table A.1. Overview of research on nonhuman primate personality. The table includes the species, sample size, age category, setting, method of assessment used, context (natural or experimental), personality traits examined, results about consistency (temporal, contextual), and additional results (sex and age differences, method comparisons).

Species	Study	N	Age	Sex	Setting	Method	Context(s)	Personality Trait(s)	Stability	Sex difference	Age difference	Method Comparison
Assamese macaque	(Ebenau et al., 2019)	107	Infant, Juvenile, Adolescent, Adult	F, M	Wild	Rating & Coding (N)	Natural context (social and non-social)	Opportunism, Confidence, Activity, Friendliness, Gregariousness, Aggressiveness, Sociability, Vigilance	Temporal consistency	Not tested	Not tested	No
Barbary macaque	(Konečná et al., 2012)	26	Juvenile, Adult	F, M	Semi-wild	Rating & Coding (N)	Natural context (agonistic, affiliative, sexual)	Friendliness, Excitability/Activity, Confidence, Opportunism, Dominance rank	Temporal consistency (rating)	No effect	No effect	No
	(Tkaczynski et al., 2018)	27	Adult	F, M	Wild	Rating & Coding (NE)	Natural context (social and non-social); Experimental situation (playback stimuli - aggression growls and alarm barks from non-group conspecifics, novel object test)	Confidence, Friendliness, Neuroticism, Excitability, Sociability, Tactility, Boldness, Exploration	Temporal and Contextual (Experimental only) consistency	Not tested	Not tested	Yes
Black tufted-eared marmoset	(Barros, Boere, Mello, & Tomaz, 2002)	7	Adult	F, M	Captive	Coding (E)	Experimental situation (taxidermized potential predators, 15 cm purple bear-like stuffed toy)	Emotional reaction / Alertness	Not tested	No effect	Not tested	No
Bonnet Macaque	(Ramakrishnan, Coss,	101	Juvenile, Subadult	F, M	Wild	Coding (E)	Experimental situation (snake tests)	Behavioural responses to snake	Not tested	Not tested	Effect	No

	Schank, Dharawat, & Kim, 2005)											
Bonobo	(Garai, Weiss, Arnaud, & Furuichi, 2016)	16	Adolescent, Adult	F, M	Wild	Rating & Coding (N)	Natural context (grooming, resting, feeding, or moving)	Unemotionality, Friendliness, Aggressiveness, Irritability, Activity, Grooming, Playfulness, Intraversion, Dominance rank	Not tested	Effect	Effect	No
	(Martin & Suarez, 2017)	24	Adolescent, Adult	F, M	Captive	Coding (N)	Natural context (social and non-social)	Agreeableness, Openness	Unclear	Effect	Effect	No
	(Staes, Eens, Weiss, & Stevens, 2016)	154	Infant, Juvenile, Adolescent, Adult	F, M	Captive	Rating & Coding (NE)	Natural context (social and non-social); Experimental situation (predator, novel food, puzzle feeder)	Assertiveness, Conscientiousness, Openness, Attentiveness, Agreeableness, Extraversion, Dominance rank	Temporal consistency	Effect	Effect	No
	(Weiss et al., 2015)	154	Juvenile, Adolescent, Adult	F, M	Captive	Rating	NA	Assertiveness, Conscientiousness, Openness, Agreeableness, Attentiveness, Extraversion	Temporal consistency	Not tested	Not tested	No
Bornean orangutan	(Russon, Kuncoro, Ferisa, & Putri Handayani, 2010)	43	Juvenile, Adolescent	F, M	Semi-wild & Wild	Coding (N)	Natural context (social and non-social)	Innovation	Not tested	No effect	No effect	No
Brown capuchin	(Morton et al., 2013)	127	Juvenile, Adult	F, M	Captive	Rating & Coding (NE)	Natural context (social and non-social); Experimental situation (free access to cognitive task)	Assertiveness, Openness, Neuroticism, Sociability, Attentiveness	Not tested	Not tested	Not tested	No
Capuchin	(Ferreira et al., 2016)	123	Adult	F, M	Semi-wild	Coding (N)	Natural context (social and non-social)	Restless, Self-narcotizing/fear, Self-protection, Stereotyped, Help-seek	Not tested	No effect	Not tested	No

	(Lefevre et al., 2014)	64	Juvenile, Adult	F, M	Captive	Rating	NA	Assertiveness	Not tested	Not tested	Not tested	No
	(Uher & Visalberghi, 2016)	150	Juvenile, Adult	F, M	Captive	Rating & Coding (NE)	Natural context (pre-feeding, afternoon); Experimental situation (14 tests)	20 constructs (e.g., Aggressiveness, Anxiousness, Gregariousness, Playfulness)	Temporal consistency	Effect	Effect	Yes
Chacma baboon	(Carter et al., 2012a)	58	Subadult, Adult	F, M	Wild	Rating & Coding (E)	Experimental situation (novel food items)	Boldness	Not tested	No effect	Effect	Yes
	(Carter et al., 2012b)	57	Juvenile, Subadult, Adult	F, M	Wild	Coding (E)	Experimental situation (novel food, snake)	Boldness, Anxiety	Not tested	Effect	Effect	No
	(Carter et al., 2014)	57	Juvenile, Subadult, Adult	F, M	Wild	Coding (E)	Experimental situation (novel food, snake)	Boldness, Anxiety, Social learning	Temporal consistency	Effect	Effect	No
	(Seyfarth et al., 2012)	45	Adult	F	Wild	Coding (N)	Natural context (social and non-social)	Aloof, Loner, Nice	Temporal consistency	Not tested	Not tested	No
Chimpanzee	(Altschul et al., 2018)	538	Juvenile, Adolescent, Adult	F, M	Captive & Semi-wild	Rating	NA	Agreeableness, Conscientiousness, Dominance, Extraversion, Neuroticism, Openness	Not tested	Effect	Effect	No
	(Anestis, 2005)	up to 48	Juvenile, Adolescent	F, M	Captive	Coding (N)	Natural context (social and non-social)	Aggressive, Smart, Mellow, Playful, Affiliative, Friendly, Dominance Rank	Not tested	No effect	No effect	No
	(Anestis, 2006)	up to 48	Juvenile, Adolescent	M	Captive	Coding (N)	Natural context (social and non-social)	Aggressive, Smart, Mellow, Playful, Affiliative, Friendly, Dominance Rank	Not tested	Not tested	No effect	No
	(Bard & Gardner, 1996)	29	Infant	F, M	Captive	Rating & Coding (NE)	Natural context; Experimental situation (human interactions, object interactions)	Alertness, Responsiveness, Exploratiiveness, Extraversion, Activity	Temporal consistency (to some extent)	Effect	Effect	No
	(Botero, MacDonald, & Miller, 2013)	6	Adolescent	F, M	Wild	Coding (N)	Natural context (social)	Anxiety-related behaviour	Not tested	Not tested	Not tested	No

(Brosnan et al., 2015)	24	Adult	F, M	Captive	Rating	Experimental situation (food preference test)	Extraversion, Methodical, Reactivity/Undependability, Agreeableness, Dominance, Openness	Not tested	No effect	No effect	No
(Buirski & Plutchik, 1991)	1	NA	F	Wild	Rating	NA	Trustful, Dyscontrolled, Timid, Depressed, Distrustful, Controlled, Aggressive, Gregarious	Not tested	Not tested	Not tested	No
(Buirski, Plutchik, & Kellerman, 1978)	23	NA	F, M	Wild	Rating	NA	Trustful, Dyscontrolled, Timid, Depressed, Distrustful, Controlled, Aggressive, Gregarious	Not tested	Effect	Not tested	No
(Clay et al., 2015)	46	Infant to Adult	F, M	Captive	Rating & Coding (N)	Natural context (social and non-social)	Dominance, Extraversion, Conscientiousness, Agreeableness	Not tested	Not tested	Not tested	No
(Clay, Bard, & Bloomsmith, 2018)	35	Infant to Adult	F, M	Captive	Rating & Coding (N)	Natural context (social and non-social)	Dominance, Extraversion, Conscientiousness, Agreeableness	Not tested	Effect	Not tested	No
(Crawford, 1938)	9	NI	NI	Captive	Rating	NA	NI	NI	NI	NI	No
(Dutton, 2008)	75	Infant, Juvenile, Adolescent, Adult	F, M	Captive	Rating	NA	Agreeableness, Dominance, Neuroticism, Extraversion, Intellect	Temporal consistency	Effect	Effect	No
(Dutton, Clark, & Dickins, 1997)	24	Infant, Juvenile, Adolescent, Adult	F, M	Captive	Rating	NA	Dominance, Sociability, Machiavellianism, Anxiety	Not tested	Not tested	Not tested	No
(Faughn et al., 2015)	54	Infant, Juvenile, Adolescent, Adult	F, M	Captive	Rating	NA	Social responsiveness, Dominance rank	Not tested	No effect	No effect	No
(Freeman et al., 2013)	99	Adolescent, Adult	F, M	Captive	Rating & Coding (N)	Natural context (social and non-social)	Extraversion, Methodical, Reactivity/Undependability, Agreeableness, Dominance, Openness	Temporal consistency	Not tested	Not tested	Yes

(Haslerud, 1938)	12	Child, Adult	NI	Captive	Coding (E)	Experimental situation (animate objects: tortoise, snake, alligator, fire, and ball; inanimate objects paired according to size with the animate ones)	Fearfulness	NI	NI	No effect	No
(Hebb, 1949)	30	NI	NI	Captive	Coding (E)	Experimental situation (presence of men, inanimate objects)	Friendly, Aggressive, Quasi-aggressive, Avoidance, Unresponsiveness, Apparently-friendly,	NI	NI	NI	No
(Herrelko, Vick, & Buchanan-Smith, 2012)	11	Adolescent, Adult	F, M	Captive	Rating & Coding (E)	Experimental situation (test related to cognitive tasks)	Dominance, Extraversion, Conscientiousness, Agreeableness, Neuroticism, Openness	Not tested	No effect	Not tested	No
(Hopper et al., 2014)	36	Adult	F, M	Captive	Rating & Coding (E)	Experimental situation (novel foraging puzzles)	Reactivity/Undependability, Dominance, Extraversion, Openness, Agreeableness, Methodical	Not tested	No effect	No effect	No
(IJzendoorn, Bard, Bakermans-Kranenburg, & Ivan, 2009)	46	Infant	F, M	Captive	Rating & Coding (E)	Experimental situation (the Infant Behavior Record from the BSID, Strange Situation Procedure)	Attachment Development	Not tested	Not tested	Not tested	No
(King & Figueredo, 1997)	100	Juvenile to Adult	F, M	Captive	Rating	NA	Dominance, Surgency, Dependability, Agreeableness, Emotionality, Openness	Not tested	Not tested	Not tested	No
(King & Landau, 2003)	128	Juvenile to Adult	F, M	Captive	Rating & Coding (N)	Natural context (agonistic, submissive, affinitive, orientation to public, solitary)	Dominance, Surgency, Dependability, Agreeableness, Emotionality, Openness	Not tested	No effect	No effect	No
(King et al., 2005)	117	Infant, Juvenile, Adolescent, Adult	F, M	Captive & Semi-wild	Rating	NA	Dominance, Surgency, Dependability, Agreeableness, Emotionality, Openness	Not tested	Effect	Effect	No

(Koski, 2011b)	75	Adolescent, Adult	F, M	Captive	Coding (N)	Natural context (social and non-social)	Sociability, Positive affect, Equitability, Anxiety, Activity	Temporal consistency	Effect	Not tested	No
(Kutsukake, 2003)	18	Adult	F, M	Wild	Coding (N)	Natural context (affiliative, aggressive)	Anxiety-related behaviour, Dominance rank, Dyadic association	Not tested	Effect	Not tested	No
(Kutsukake et al., 2012)	14	Adult	M	Captive	Coding (E)	Experimental situation (playback stimuli: chimpanzee vocalization, jungle crow, no sound)	Neuroticism	Temporal and Contextual consistency	Not tested	Not tested	No
(Latzman, Hopkins, Keebaugh, & Young, 2014)	174	Subadult, Adult	F, M	Captive	Rating	NA	Dominance, Conscientiousness, Extraversion, Agreeableness, Intellect	Not tested	No effect	Not tested	No
(Latzman, Hecht, Freeman, Schapiro, & Hopkins, 2015)	107	Juvenile, Adolescent, Adult	F, M	Captive	Rating	NA	Reactivity/Unpredictability, Dominance, Extraversion, Openness, Agreeableness	Not tested	Effect	Effect	No
(Latzman, Freeman, et al., 2015)	238	Subadult, Adult	F, M	Captive	Rating	NA	Reactivity/Unpredictability, Dominance, Extraversion, Openness, Agreeableness	Not tested	Effect	Effect	No
(Lehmann & Boesch, 2009)	NI	Adolescent, Adult	F, M	Wild	Coding (N)	Natural context (social and non-social)	Sociality	Not tested	Effect	Not tested	No
(Lilienfeld, Gershon, Duke, Marino, & de Waal, 1999)	34	Juvenile, Adolescent, Adult	F, M	Captive	Rating & Coding (N)	Natural context (social and non-social)	Psychopathy, Extraversion, Agreeableness, Neuroticism	Not tested	Effect	No effect	No
(Martin, 2005)	43	Adolescent, young Adult, Adult	F, M	Captive	Rating	NA	5 components but not labelled	Not tested	Effect	No effect	No

(Massen & Koski, 2014)	14 to 15, 15 to 22	Adolescent, Adult	F, M	Captive	Coding (NE)	Natural context (social and non-social; see Koski (2011)); Experimental situation (see Massen et al (2013))	Sociability, Positive affect, Equitability, Anxiety, Boldness, Exploration tendency-persistence, Relationship quality	Not tested	No effect	No effect	No
(Massen et al., 2013)	29	Adolescent, Adult	F, M	Captive	Coding (E)	Experimental situation (10 tests)	Exploration tendency, Persistence, Boldness, Tool-Orientation	Temporal and Contextual consistency	No effect	Effect	No
(Murray, 1998)	59	Infant, Juvenile, young Adult, mature Adult	F, M	Captive	Rating	NA	Confident/Apprehensive, Sociable/Solitary, Excitable/Slow	Not tested	Effect	Effect	No
(Pederson et al., 2005)	49	Juvenile, Adult	F, M	Captive	Rating & Coding (N)	Natural context (agonistic, submissive, affinitive, orientation to public, solitary)	Extraversion, Dominance, Dependability, Agreeableness, Emotionality, Openness	Not tested	Effect	Not tested	No
(van Hooff, 1973)	25	Infant, Juvenile, Adolescent, Adult	F, M	Captive	Coding (N)	Natural context (social)	Social behaviours	Not tested	Not tested	Not tested	No
(Leeuwen, Cronin, & Haun, 2018)	89	Infant, Juvenile, Adolescent, Adult	F, M	Semi-wild	Coding (N)	Natural context (social and non-social)	Social Dynamics (not personality per se)	Temporal consistency	Effect	Effect	No
(Weiss et al., 2007)	379	Infant, Juvenile, Adult	F, M	Captive	Rating	NA	Dominance, Extraversion, Conscientiousness, Agreeableness, Neuroticism, Openness	Not tested	Effect	Effect	No
(Weiss et al., 2009)	146	Infant, Juvenile, Adolescent, Adult	F, M	Captive & Semi-wild	Rating	NA	Dominance, Extraversion, Conscientiousness, Agreeableness, Neuroticism, Openness	Not tested	No effect	Effect	No
(Weiss et al., 2017)	128	Infant to adult	F, M	Wild	Rating	NA	Dominance, Extraversion, Conscientiousness, Agreeableness, Neuroticism, Openness	Temporal consistency	Not tested	Not tested	No

	(Yerkes, 1940)	NI	NI	F, M	Captive	Coding (N)	Natural context (social)	Sociability	NI	NI	NI	No
	(Yerkes & Yerkes, 1936)	NI	Infant, Children, Adult	NI	Captive	Coding (E)	Experimental situation (shuttlecock, rubber dog, rubber tube, live tortoise, glass snake)	Fearfulness	NI	NI	Effect	No
Chimpanzee (and Human)	(King et al., 2008)	202	Infant, Juvenile, Adolescent, Adult	F, M	Captive	Rating	NA	Dominance (chimp only), Extraversion, Agreeableness, Conscientiousness, Neuroticism, Openness to Experience	Temporal consistency	Effect	Effect	No
	(Marrus et al., 2011)	29 (20)	Juvenile, Adolescent, Adult	F, M	Captive & Semi-wild	Rating & Coding (E)	Experimental situation (object handling)	Social behaviours	Not tested	No effect	No effect	No
Chimpanzee, Bonobo, Orangutan (and Human)	(Herrmann et al., 2011)	24, 24, 24 (105, 14)	Adolescent, Adult	F, M	Captive & Semi-wild	Coding (E)	Experimental situation (novel object, novel food, people)	Shyness-Boldness	Not tested	No effect	No effect	No
Chimpanzee, Bonobo, Orangutan, Gorilla	(Uher & Asendorpf, 2008)	5, 5, 5, 5	Adolescent, Adult	F, M	Captive	Rating & Coding (NE)	Natural context (pre-feeding, afternoon); Experimental situation (14 tests)	17 personality traits (e.g., Anxiousness, Curiosity, Impulsiveness)	Temporal and Contextual consistency	Not tested	Not tested	Yes
	(Uher et al., 2008)	5, 5, 5, 5	Adolescent, Adult	F, M	Captive	Coding (NE)	Natural context (pre-feeding, afternoon); Experimental situation (14 tests)	Aggressiveness, Arousability, Anxiousness, Curiosity, Friendliness to humans, Impulsivity, Persistency, Physical activity	Temporal and Contextual consistency	Not tested	Not tested	No
Chimpanzee, Orangutan	(Weiss & King, 2015)	174, 202	Infant, Juvenile, Adolescent, Adult	F, M	Captive	Rating	NA	Extraversion, Dominance, Neuroticism, Agreeableness, Intellect	Not tested	Effect	Effect	No
	(Weiss et al., 2012)	357, 174	Infant, Juvenile, Adolescent, Adult	F, M	Captive & Semi-wild	Rating	NA	Dominance, Extraversion, Conscientiousness, Agreeableness,	Not tested	Not tested	Not tested	No

								Neuroticism, Intellect/Openness				
Chimpanzee, Orangutan, Rhesus macaque	(Weiss, Adams, & Johnson, 2011)	100, 152, 111	Infant, Juvenile, Adolescent, Adult	F, M	Captive & Semi-wild	Rating	NA	Dominance, Extraversion, Conscientiousness, Agreeableness, Neuroticism, Openness, Intellect, Confidence, Openness, Friendliness, Anxiety, Activity	Not tested	Not tested	Not tested	No
Common marmoset	(Burkart, Strasser, & Foglia, 2009)	38	Adult	F, M	Captive	Coding (E)	Experimental situation (video stimuli, problem-solving task, baited Perspex box, maze)	Social learning & Innovation	Not tested	Effect	No effect	No
	(Inoue-Murayama et al., 2018)	77	Infant, Juvenile, Adult	F, M	Captive	Rating	NA	Dominance, Sociability, Neuroticism	Not tested	Not tested	Not tested	No
	(Iwanicki & Lehmann, 2015)	63	Juvenile, Adult	F, M	Semi-wild	Rating & Coding (N)	Natural context (social and non-social)	Extraversion, Agreeableness, Neuroticism, Openness, Conscientiousness	Not tested	Not tested	Not tested	No
	(Koski & Burkart, 2015)	17	Adult	F, M	Captive	Coding (E)	Experimental situation (8 tests)	Boldness, Exploratory tendency, Persistence	Temporal and Contextual consistency	No effect	Not tested	No
	(Koski et al., 2017)	100	Infant, Juvenile, Adolescent, Adult	F, M	Captive	Rating	NA	Conscientiousness, Agreeableness, Inquisitiveness, Assertiveness, Patience	Not tested	Effect	Effect	No
	(Šlipogor et al., 2016)	21	Adolescent, Adult	F, M	Captive	Coding (E)	Experimental situation (general activity, novel food, novel object, predator and foraging under risk)	Boldness-Shyness, Stress-Activity, Exploration-Avoidance	Temporal and Contextual consistency	No effect	No effect	No
	(Tomassetti et al., 2019)	16	Adult	F, M	Captive	Coding (NE)	Natural context (social); Experimental situation (10 tests based on Uher et al 2008)	13 personality traits (e.g., aggressiveness, anxiousness, arousability, curiosity)	Temporal consistency	No effect	No effect	No

Crab-eating macaque	(Uher, Werner, et al., 2013)	104	Subadult, Adult	F, M	Captive	Rating & Coding (N)	Natural context (social and non-social)	18 macaque-specific personality constructs.	Temporal consistency	Effect	Effect	Yes
Crab-eating macaque, Tufted capuchin	(Vitale, Visalberghi, & De Lillo, 1991)	4, 6	Juvenile, Adult	F, M	Captive	Coding (E)	Experimental situation (rubber snake)	Behavioural responses to snake	Temporal consistency	Not tested	Not tested	No
Crested macaque	(Neumann et al., 2013)	37	Adult	M	Wild	Coding (NE)	Natural context (social); Experimental situation (playback experiments - dog bark, donkey bray)	Anxiety, Connectedness, Sociability, Aggressiveness, Boldness	Temporal consistency	Not tested	Not tested	Yes
Crested macaque, Barbary macaque, Common squirrel monkey	(Baker et al., 2015)	64, 62, 69	Juvenile, Adult	F, M	Captive	Rating & Coding (NE)	Natural context (social and non-social); Experimental situation (novel object)	Dominance, Sociability, Emotionality, Human-Animal Sociability, Cautiousness	Not tested	Effect	Effect	No
Cynomolgus macaque	(Kaplan, Manuck, Fontenot, & Mann, 2002)	46	Adult	F, M	Captive	Coding (N)	Natural context (social and non-social)	Dominance	Not tested	No effect	Not tested	No
Cynomolgus macaque, lion-tailed macaque	(Clarke & Lindburg, 1993)	5, 5	NA	M	Captive	Coding (NE)	Natural context (social); Experimental situation (unfamiliar cage)	Boldness, Curiosity, Instrumental	NI	NI	NI	No
Geoffroy's marmoset	(Hankerson & Caine, 2004)	11	Adolescent, Adult	F, M	Captive	Coding (E)	Experimental situation (freeze-dried rattlesnake, piece of cloth)	Vigilance Behaviours	Not tested	Not tested	Not tested	No
Gorilla	(Gold & Maple, 1994)	298	Infant, Juvenile, Adolescent, Adult	F, M	Captive	Rating	NA	Extroverted, Dominant, Fearful, Understanding	Not tested	Not tested	Not tested	No
	(Kuhar et al., 2006)	119	Infant, Juvenile, Adolescent, Adult	M	Captive	Rating & Coding (N)	Natural context (social and non-social)	Extraversion, Dominance, Fearful, Understanding	Not tested	Not tested	Effect	No

Grey mouse lemur	(Dammhahn, 2012)	117	NA	F, M	Wild	Coding (E)	Experimental situation (open field, novel object)	Boldness, Exploration, Activity	Temporal consistency	Effect	Effect	No
	(Dammhahn & Almeling, 2012)	36	Juvenile, Adult	F, M	Wild	Coding (E)	Experimental situation (risk-sensitive foraging test, novel object test)	Boldness, Explorativeness	Temporal and Contextual consistency	No effect	No effect	No
Grivet monkey	(Fairbanks et al., 1999)	22	Infant to Adult	F, M	Wild	Coding (E)	Experimental situation (trapped - re-trapped)	Impulsivity	Not tested	No effect	No effect	No
Hanuman langur	(Konečná et al., 2008)	27	Subadult, Adult	M	Semi-wild	Rating & Coding (N)	Natural context (social and non-social)	Agreeableness, Confidence, and Extraversion, Dominance rank	Not tested	Not tested	Not tested	No
Japanese macaque	(French, 1981)	3	Infant and Mother	NI	Captive	Coding (N)	Natural context (play interaction)	Playfulness, Maternal style	NI	NI	NI	No
	(Troisi et al., 1991)	7	Infant and mother	F	Captive	Coding (N)	Natural context (pre and post-birth)	Anxiety, Maternal style, Dominance	Not tested	Not tested	Not tested	No
Japanese macaque, Gothic-arch squirrel monkeys, Roman-arch squirrel monkeys, Hamadryas baboon	(Martau, Caine, & Candland, 1985)	14, 6, 5, 8	Infant, Juvenile, Adolescent, Adult	F, M	Captive	Rating	NA	NI	Temporal consistency	NI	NI	No
Japanese macaque, Rhesus macaque	(Bardi et al., 2001)	7, 3	Infant and mother	F, M	Captive	Coding (N)	Natural context (social)	Rejection, Protectiveness, Independence, Maternal style	Temporal consistency (to some extent)	Not tested	Not tested	No
Lion-tailed macaque	(Rouff, Sussman, & Strube, 2005)	52	NA	F, M	Captive	Coding (NE)	Natural context (social and non-social); Experimental situation (novel object, persistent test)	Extraversion-like behaviours, Agonistic behaviours, Bold/Cautious behaviours	Not tested	Effect	Not tested	No

	(Pavani, Maestripieri, Schino, Turillazzi, & Scucchi, 1991)	8	Adult	F, M	Captive	Coding (N)	Natural context (social and non-social)	Anxiety-related behaviour, Social proximity	Not tested	Effect	No effect	No
	(Wergård, Westlund, Spångberg, Fredlund, & Forkman, 2016)	34	Adult	F, M	Captive	Rating & Coding (N)	Natural context (social rank assessment)	Emotionality, Activity, Sociability, Tolerance, Social rank	Not tested	Not tested	Not tested	No
Mountain gorilla	(Eckardt et al., 2015)	116	Juvenile, Adolescent, Adult	F, M	Wild	Rating & Coding (N)	Natural context (social and non-social)	Dominance, Openness, Sociability, Proto-Agreeableness	Not tested	Effect	Effect	No
Olive baboon	(Buirski, Kellerman, Plutchik, Weininger, & Buirski, 1973)	7	Infant, Juvenile, Adult	F, M	Wild	Coding (N)	Natural context (social and non-social)	Incorporation, Orientation, Protection, Deprivation, Rejection, Exploration, Destruction, Dominance rank, Social behaviour	Not tested	Not tested	Not tested	No
	(Castles et al., 1999)	10	older Subadult, Adult	F	Wild	Coding (N)	Natural context (drink, eat, forage, travel search, travel, immobile, social, solitary)	Anxiety	Not tested	Not tested	Not tested	No
	(Sapolsky & Ray, 1989)	NI	NI	M	Wild	Coding (N)	NI	Dominance, Aggressiveness	NI	NI	NI	No
Olive baboon, Yellow baboon	(Johnson et al., 2015)	578	Adult	F, M	Captive	Coding (NE)	Natural context (social and non-social); Experimental situation (novel object, apparent novel social partner using mirror test)	Boldness, Engagement with object	Not tested	Effect	Effect	No
Pig-tailed macaque	(Caine, Earle, & Reite, 1983)	10	Adolescent	NI	Captive	Rating	NA	Confidence, Opportunistic, Effective, Insecure, Sociability, Dominance	NI	NI	NI	No
	(Reite & Short, 1980)	21	Infant	F, M	Captive	Coding (N)	Natural context (social and non-social)	Activity	Temporal consistency	NI	NI	No

	(Sussman & Ha, 2011)	167	Infant	F, M	Captive	Coding (E)	Experimental situation (daily health measure test - handling, novelty probe - tester wears different outfit)	Reactivity & Boldness	Temporal and Contextual consistency	No effect	Effect	No
	(Sussman, Mates, Ha, Bentson, & Crockett, 2014)	293	Adult	F, M	Captive	Coding (E)	Experimental situation (rapid Assessment of Temperament and Reactivity test)	Sociability toward humans, Cautiousness, Aggressiveness, Fearfulness	Temporal consistency	Effect	Effect	No
Pig-tailed macaque, Bonnet macaque	(Laudenslager & Boccia, 1996)	NI	Infant, Adult	F, M	Captive	Coding (E)	Experimental situation (social separation, social conflict, temporal restriction, spatial restriction, social intruder)	Social affiliation	NI	NI	NI	No
Pig-tailed macaque, Long-tailed macaque, Yellow baboon	(Heath-Lange, Ha, & Sackett, 1999)	7, 3, 4	Infant	M	Captive	Coding (E)	Experimental situation (response to capture)	Reactivity	No temporal consistency found	Not tested	Effect	No
Pig-tailed macaque, Rhesus macaque	(Westergaard, Mehlman, Westergaard, Suomi, & Higley, 1999)	30, 31	Juvenile, Adolescent	F	Captive	Coding (N)	Natural context (social and non-social)	Impulsivity, Aggression	Not tested	Not tested	Not tested	No
Red-chested tamarin, Saddle-back tamarin, Cotton-top tamarin	(Box, Röhrhuber, & Smith, 1995)	24	Adult	F, M	Captive	Coding (E)	Experimental situation (foraging task)	Foraging-related behaviours	Not tested	Effect	Not tested	No
Rhesus macaque	(Altschul, Robinson, Coleman, Capitanio, & Wilson, 2019)	109	Infant, Juvenile, Adult	F, M	Captive	Rating & Coding (N)	Natural context (social and non-social)	Confidence, Openness, Assertiveness, Friendliness, Activity, Anxiety, Dominance rank	Not tested	Not tested	Effect	No

(Bolig, Price, O'Neill, & Suomi, 1992)	22	Infant, Juvenile, Adolescent, Adult	F, M	Captive	Rating	NA	Reactivity, 10 other personality traits (NI)	NI	NI	NI	No
(Brent et al., 2014)	108	Adult	F	Captive & Semi-wild	Coding (N)	Natural context (social and non-social)	Meek, Bold, Aggressive, Passive, Loner, Nervous, Dominance rank	Temporal consistency	Not tested	Effect	No
(Capitanio, 1984)	12	Infant	F, M	Captive	Coding (E)	Experimental situation (animate nonconspecific, inanimate mother-substitute)	Social ability	NI	NI	NI	No
(Capitanio, 1985)	12	Infant	F, M	Captive	Coding (NE)	Natural context (social); Experimental situation (animate nonconspecific, inanimate mother-substitute)	Social ability	NI	NI	NI	No
(Capitanio, 1999)	45	Adult	M	Captive	Rating & Coding (NE)	Natural context (social and non-social); Experimental situation (responsiveness to a threatening human, stable and unstable group, responses to videotapes of social behaviour, social dyad study)	Sociability, Confidence, Equability, Excitability	Not tested	Not tested	Not tested	No
(Capitanio, 2002)	12	Adult	M	Captive	Rating & Coding (E)	Experimental situation (video stimuli: Agression, Affiliation, Non-social)	Sociability, Confidence, Equability, Excitability	Not tested	Not tested	Not tested	No
(Capitanio, Mendoza, & Baroncelli, 1999)	42	Adult	M	Captive	Rating	NA	Sociability, Confidence, Equability, Excitability	Not tested	Not tested	Not tested	No
(Capitanio, Mendoza, & Bentson, 2004)	42	Adult	M	Captive	Rating	NA	Sociability, Confidence, Equability, Excitability	Not tested	Not tested	Not tested	No

(Chamove, Eysenck, & Harlow, 1972)	168	Juvenile	NI	Captive	Coding (N)	Natural context (stable quadrad peer groups, in newly-formed dyads with infant, juvenile, and adult stimulus monkeys; in similar triads with the stimulus animal plus a familiar cage-mate)	Affiliative, Hostile, Fearful	NI	NI	NI	No
(Clarke & Snipes, 1998)	48	Infant	F, M	Captive	Rating	NA	Fearfulness, Excitability, Activity, Attentiveness, Cautiousness	No temporal consistency found	Effect	NI	No
(Davidson, Kalin, & Shelton, 1993)	9	NI	NI	Captive	Coding (E)	Experimental situation (medical experiment)	Behavioral inhibition	NI	NI	NI	No
(Freedman & Rosvold, 1962)	7	Adult	F, M	Captive	Coding (E)	Experimental situation (standard laboratory conditions)	Aggressiveness, Anxiety, Sexual behaviour	NI	Not tested	Not tested	No
(Hinde et al., 2015)	108	Infant and Mother	F	Captive	Rating (E)	Experimental situation	Nervousness, Confidence	Not tested	Not tested	Not tested	No
(Kalin, Shelton, Rickman, & Davidson, 1998)	28	Infant and Mother	F, M	Captive	Coding (E)	Experimental situation (human intruder paradigm)	Freezing behaviour	Temporal consistency	No effect	Effect	No
(Kalin, Shelton, Davidson, & Kelley, 2001)	17	Adolescent	F, M	Captive	Coding (E)	Experimental situation (human intruder paradigm, social threat, snake fear test)	Anxiousness	Not tested	Not tested	Not tested	No
(Kalin, Shelton, Fox, Oakes, & Davidson, 2005)	25	Adolescent, Adult	M	Captive	Coding (E)	Experimental situation (human intruder paradigm)	Anxiousness	Not tested	Not tested	Not tested	No
(Kalin, Shelton, & Davidson, 2007)	12	Adolescent	M	Captive	Coding (E)	Experimental situation (human intruder paradigm, modified human intruder paradigm)	Anxiousness	Not tested	Not tested	Not tested	No

(Locke, Locke, et al., 1964)	12	Infant	F, M	Captive	Coding (N)	Natural context (play interaction)	Approach-avoidance factor, Avoidance-approach factor	Temporal consistency but no Contextual consistency	No effect	Effect	No
(Locke, Morgan, & Zimmermann, 1964)	12	Infant	F, M	Captive	Coding (N)	Natural context (social)	Social behaviours	Not tested	Not tested	Not tested	No
(Maestripieri, 2000)	10	Adult	F	Captive	Coding (N)	Natural context (birth season, mating season)	Emotionality	Temporal consistency	Not tested	Not tested	No
(McCowan et al., 2011)	60	Unclear	F, M	Captive	Coding (N)	Natural context (social and non-social)	Bold/Confident/Direct, Unpredictable/Impulsive/Reckless/Aggressive/Excitable/Active/Vigilant, Tolerant/Calm/Gentle/Understanding/Popular, Affiliative/Warm	Not tested	Not tested	Not tested	No
(Mineka, Keir, & Price, 1980)	20	Adult	F, M	Captive	Coding (E)	Experimental situation (real snake, range of snake-like objects)	Fearfulness	Temporal consistency	NI	NI	No
(Robinson et al., 2018)	85	Infant, Juvenile, Adolescent, Adult	F, M	Captive	Rating & Coding (N)	Natural context (social and non-social)	Dominance, Confidence, Openness, Anxiety, and Friendliness, Dominance	Not tested	No effect	Not tested	No
(Rogers, Shelton, Shelledy, Garcia, & Kalin, 2008)	285	Juvenile	F, M	Captive	Coding (E)	Experimental situation (modified human intruder paradigm)	Anxiety	Not tested	Effect	Effect	No
(Rommeck, Capitanio, Strand, & McCowan, 2011)	32	Infant	F, M	Captive	Rating & Coding (E)	Experimental situation (holding cage observations, human intruder)	Vigilant, Gentle, Confident, Nervous, Early social experience	No temporal consistency found	Not tested	Not tested	No

(Schneider, Moore, Suomi, & Champoux, 1991)	23	Infant	F, M	Captive	Rating & Coding (E)	Experimental situation (battery of developmental tests)	Fearfulness	Unclear	No effect	Not tested	No
(Spencer-Booth & Hinde, 1969)	16	Infant	F, M	Captive	Coding (E)	Experimental situation (battery of tests; in mildly disturbing or frustrating situations)	Behavioural responses to strange objects	Temporal and Contextual consistency	Not tested	Not tested	No
(Stevenson-Hinde & Zunz, 1978)	46	Infant, Adolescent, Adult	F, M	Captive	Rating	NA	Confident-fearful, Active-slow, Sociable-solitary	Temporal consistency	Effect	Effect	No
(Stevenson-Hinde, Stillwell-Barnes, & Zunz, 1980a)	31	Infant	F, M	Captive	Coding (E)	Experimental situation (ball test, food test, mask test, mirror test, slide test, smartie test)	Behavioural measure (e.g., Activity, Approach, Gaze)	Temporal and Contextual consistency	Effect	Effect	No
(Stevenson-Hinde, Stillwell-Barnes, et al., 1980b)	46	Infant, Adolescent, Adult	F, M	Captive	Rating	NA	Confident, Excitable, Sociable	Temporal consistency	Effect	Effect	No
(Stevenson-Hinde, Zunz, & Stillwell-Barnes, 1980)	25	Infant and Mother	F, M	Captive	Coding (E)	Experimental situation (strange situation)	Confidence, Fearfulness	Temporal and Contextual consistency	Effect	NI	No
(Suarez-Jimenez et al., 2013)	26	Infant	F, M	Captive	Coding (E)	Experimental situation (Brazelton neonatal assessment, human intruder paradigm, human intruder-startle paradigm)	State control, Startle, Fear, Restrain, Isolation locomotion, Isolation vocalizations, Dominance status of mother	Not tested	No effect	Effect	No

	(Sullivan, Mendoza, & Capitanio, 2011)	179	Infant	F, M	Captive	Rating & Coding (E)	Experimental situation (separation and relocation, recognition memory, responses to social stimuli, contact with novel objects, and behavioural response to graded conditions of challenge)	Confident, Gentle, Vigilant, Nervous	Not tested	Effect	Not tested	No
	(Suomi et al., 1996)	8	Juvenile, Adult	F, M	Captive	Coding (N)	Natural context (active manipulation of their home-cage environment, social, non-social)	31 categories of behaviour (e.g., locomotion, social contact, aggression, ventral contact); Social; Explorative	Temporal consistency	Effect	Effect	No
	(von Borell et al., 2016)	24	Infant to Subadult	F, M	Semi-wild	Coding (N)	Natural context (social)	Aggression, Sociability, Fearfulness	Temporal consistency	Effect	Effect	No
	(Weinstein & Capitanio, 2012)	29	Adolescent	F, M	Captive	Rating & Coding (NE)	Natural context (social and non-social); Experimental situation (battery of tests)	Adaptability, Confidence, Equability, Frienship, Dominance rank	Temporal consistency	Effect	Not tested	No
	(Weiss, Adams, Widdig, et al., 2011)	125	Infant, Juvenile, Adolescent, Adult	F, M	Semi-wild	Rating	NA	Confidence, Openness, Dominance, Friendliness, Activity, Anxiety	Temporal consistency	Not tested	Not tested	No
	(Williamson et al., 2003)	85	Infant	F, M	Captive	Coding (NE)	Natural context (free play); Experimental situation (remote-controlled car, human intruder, novel fruit)	Anxiousness	Not tested	Not tested	Not tested	No
Rhesus macaque, Japanese macaque, Assamese macaque, Barbary macaque, Tonkean macaque, Crested macaque	(Adams et al., 2015)	125, 74, 60, 74, 46, 53	Infant, Juvenile, Adult, senior Adult	F, M	Captive & Semi-wild & Wild	Rating	Na	Dominance, Confidence, Openness, Friendliness, Anxiety, Activity, Sociality	Not tested	Effect	Not tested	No

Rhesus macaque, Long-tailed macaque, Pig-tailed macaque	(Sussman, Ha, Bentson, & Crockett, 2013)	129, 214, 556	Infant, Juvenile, Adult	F, M	Captive	Coding (E)	Experimental situation (Rapid Assessment of Temperament and Reactivity test)	Sociability toward humans, Cautiousness, Aggressiveness, Fearfulness	Not tested	Effect	Effect	No
Small-Eared Bushbaby	(Watson & Ward, 1996)	45	Adolescent, Adult	F, M	Captive	Coding (E)	Experimental situation (open field test, novel stimuli)	Activity, Boldness, Curiosity, Propensity to flee novel situations	Not tested	No effect	Effect	No
Stump-tailed macaque	(Mondragon-Ceballos & Santillán-Doherty, 1994)	27	Infant, Juvenile, Adolescent, Adult	F, M	Captive	Rating & Coding (N)	Natural context (social and non-social)	Confident, Sociable, Excitable, Somatotype, Dominance rank	Not tested	Effect	Effect	No
	(Santillán-Doherty, Muñóz-Delgado, Arenas, Márquez, & Cortés, 2006)	22	Adolescent, Adult	F, M	Captive	Rating & Coding (E)	Experimental situation (risk-taking test, curiosity test)	Novelty-seeking, Risk-taking	Not tested	Not tested	Not tested	No
Stump-tailed macaque (and Zebra Finch)	(Figueredo et al., 1995)	13	Adult	F, M	Captive	Rating	NA	Confident, Excitable, Sociable	Temporal consistency	Not tested	Not tested	No
Stump-tailed macaque, Spider monkey	(Santillán-Doherty et al., 2010)	22, 7	Adolescent, Adult	F, M	Captive	Rating & Coding (E)	Experimental situation (risk-taking test, curiosity test)	Novelty-seeking, Risk-taking, Curiosity	Not tested	Effect	Effect	No
Sumatran orangutan, Bornean orangutan	(Damerius, Graber, et al., 2017)	61	Infant, Juvenile, Adolescent, Adult	F, M	Semi-wild	Coding (E)	Experimental situation (novelty response and exploration tasks, physical cognitive tasks)	Curiosity/Explorativeness, Cognitive performance	Not tested	No effect	No effect	No

	(Damerius, Forss, et al., 2017)	103	Infant, Juvenile, Adolescent, Adult	F, M	Captive	Coding (E)	Experimental situation (novelty response and exploration tasks, physical cognitive tasks, reactions towards an unfamiliar human and novel food)	Curiosity/Explorativeness, Cognitive performance, Human Orientation	Not tested	No effect	No effect	No
	(Forss et al., 2015)	14, 28	Juvenile, Adolescent, Adult	F, M	Captive & Wild	Coding (E)	Experimental situation (novel object)	Curiosity/Explorativeness	Not tested	Not tested	Not tested	No
	(Forss et al., 2016)	33	Juvenile, Adolescent, Adult	F, M	Captive	Coding (E)	Experimental situation (physical cognition task - honey tool-task, detour reaching task, reversal-learning task; novel food, novel toy)	Explorativeness	Not tested	Effect	Effect	No
	(Schuppli, Forss, et al., 2016)	26	Infant, Juvenile	F, M	Wild	Coding (N)	Natural context (feeding, social)	Developmental measures regarding feeding behaviour (exploration)	Not tested	No effect	No effect	No
	(Schuppli et al., 2017)	21, 11	Infant, Juvenile and Mothers	F, M	Wild	Coding (N)	Natural context (social and non-social)	Explorativeness, Social learning	Not tested	No effect	Effect	No
	(Weiss et al., 2006)	152	Infant, Juvenile, Adolescent, Adult	F, M	Captive	Rating	NA	Extraversion, Dominance, Neuroticism, Agreeableness, Intellect	Not tested	Not tested	Not tested	No
Sumatran orangutan, Bornean orangutan, Hybrid orangutan	(Adams et al., 2012)	184	Juvenile, Adolescent, Adult	F, M	Captive	Rating	NA	Extraversion, Dominance, Neuroticism, Agreeableness, Intellect	Not tested	Not tested	Not tested	No
	(Weiss, Adams, & King, 2011)	184	Juvenile, Adolescent, Adult	F, M	Captive	Rating	NA	Extraversion, Dominance, Neuroticism, Agreeableness, Intellect	Not tested	No effect	No effect	No
Tibetan macaque	(Pritchard, Sheeran, Gabriel, Li, & Wagner, 2014)	12	Adult	F, M	Semi-wild	Rating & Coding (N)	Natural context (social and non-social)	Insecurity, Reactivity, Boldness, Sociability, Leadership	Not tested	Not tested	Not tested	No

	(Chen et al., 2018)	21	Adult	F, M	Semi-wild	Coding (N)	Natural context (social and non-social)	Leadership, Introversion, Sociability, Solitary, Bullying, Insecurity, Nervousness, Affiliation	NI	Effect	NI	No
Tufted capuchin	(Byrne & Suomi, 1995)	17	Infant	NI	Captive	Coding (N)	Natural context (exploration environment, social interaction with group members)	Exploration, Sociability	Temporal consistency	NI	NI	No
	(Byrne & Suomi, 1998)	29	Infant	NI	Captive	Coding (N)	Natural context (exploration environment, social interaction with group members)	Exploration, Sociability	NI	NI	NI	No
	(Byrne & Suomi, 2002)	36	Juvenile	NI	Captive	Rating & Coding (N)	Natural context (social and non-social)	Aggressive, Confident, Curious, Effective, Opportunistic, Apprehensive, Fearful, Insecure, Submissive, Tense, Fearful, Play and Exploration behaviour	Not tested	NI	NI	No
	(Uher, Addessi, et al., 2013)	26	Adult	F, M	Captive	Coding (NE)	Natural context (pre-feeding, afternoon); Experimental situation (14 tests)	20 personality constructs (e.g., vigilance, anxiety, impulsivity, curiosity)	Temporal and Contextual consistency	Effect	Effect	No
Vervet monkey	(Fairbanks, 1993)	Abo ut 86	Infant, Juvenile, Adult	F, M	Captive & Wild	Coding (NE)	Natural context (habituation to a human observer); Experimental situation (novel food test, new area, come closer to a stranger adult male)	Risk-taking	Not tested	No effect	Effect	No
	(Fairbanks, 2001)	128	Adolescent, Adult	M	Captive	Coding (E)	Experimental situation (intruder challenge test)	Impulsivity, Anxiety, Dominance rank	Temporal consistency	Not tested	Effect	No
	(Fairbanks & McGuire, 1993)	83	Adult	F	Captive	Coding (NE)	Natural context (mother infant interaction); Experimental situation (novel food test)	Maternal protectiveness & Exploratiiveness	NI	NI	NI	No
	(Fairbanks, Melega, Jorgensen, Kaplan, &	138	Adolescent, Adult	M	Captive	Coding (E)	Experimental situation (intruder challenge test)	Impulsivity, Anxiety	Not tested	Not tested	Effect	No

	McGuire, 2001)											
	(Fairbanks, Jorgensen, et al., 2004)	36	Adolescent	M	Captive	Coding (E)	Experimental situation (intruder challenge test)	Impulsivity, Anxiety, Dominance rank	Temporal consistency (to some extent)	Not tested	Effect	No
	(Fairbanks, Newman, et al., 2004)	352	Adolescent, Adult	F, M	Captive	Coding (E)	Experimental situation (intruder challenge test)	Impulsivity, Aggressivity	Temporal consistency	Effect	Effect	No
	(Fairbanks et al., 2011)	503	Adolescent, Adult	F, M	Captive	Coding (E)	Experimental situation (home group novelty test)	Exploratiiveness	Temporal consistency	No effect	Effect	No
	(James et al., 2007)	18	Adolescent	M	Captive	Coding (E)	Experimental situation (intruder challenge test, home group novelty test - threatening object)	Impulsivity, Novelty seeking	Not tested	Not tested	Not tested	No
	(Jarrett et al., 2018)	28	Infant and Mother	F, M	Wild	Coding (N)	Natural context (social)	Sociability (grooming network)	No temporal consistency found	Effect	Not tested	
	(Josephs, Bonnell, Dostie, Barrett, & Henzi, 2016)	68	Infant, Juvenile, Adolescent, Adult	F, M	Wild	Coding (N)	Natural context (social and non-social)	Sociability, Dominance	Not tested	Not tested	Not tested	No
	(McFarland et al., 2017)	13 to 21	Adult	F	Wild	Coding (N)	Natural context (social and non-social)	Sociability, Aggressiveness	Not tested	Not tested	Not tested	No
	(McGuire et al., 1994)	97	Juvenile, Subadult, Adult (from 19 months to 14 years)	F, M	Captive	Rating & Coding (N)	Natural context (social and non-social)	Socially competent, Playful, Curious, Opportunistic	Temporal consistency	Effect	Effect	No
Western lowland gorilla	(Schaefer & Steklis, 2014)	8	Subadult, Adult	M	Captive	Rating & Coding (N)	Natural context (social and non-social)	Dominance, Extraversion/Agreeableness, Conscientiousness	Not tested	Not tested	Not tested	No

	(Weiss, Gartner, Gold, & Stoinski, 2013)	283	Infant, Juvenile, Adolescent, Adult	F, M	Captive	Rating	NA	Dominance, Extraversion, Neuroticism, Agreeableness	Not tested	No effect	No effect	No
	(Robbins et al., 2016)	Several	Infant, Juvenile, Adolescent, Adult	F, M	Wild	Coding (N)	Natural context (social and non-social)	Cultural traits (related to foraging, environment, social interactions, gestures, communication)	Not tested	Not tested	Not tested	No
White-faced Capuchin	(Manson & Perry, 2013)	240	Juvenile, Adult	F, M	Wild	Rating & Coding (N)	Natural context (social and non-social)	Extraversion, Openness, Neuroticism, Agreeableness, Eccentricity	Temporal consistency	Effect	Effect	No

F: female; M: male; N: naturalistic method; E: experimental method; NE: naturalistic and experimental method; NI: information not accessible; NA: not applicable

B. Appendix 2. Supplementary material for Chapter 3

Table B.1. Subject representation and colony composition. Overview of the number of subjects for each year of recording per age and sex group per colony. The number of chimpanzees that made up each colony is shown in parenthesis; Infants were not included in the study. A total of four adults/subadult chimpanzees (three from Colony 1 and one from Colony 4) and one juvenile (from Colony 1) from the 2013 recording period died prior to the 2017 recording period. Four infants (three from Colony 1 and one from Colony 4) were reported, for the 2017 recording period, born in between the two recording periods. Adults/subadults were over 8 years old, juveniles were 4-8 years old, and infants were less than 4 years old.

		2013		2017	
Age group	Sex	Colony 1	Colony 4	Colony 1	Colony 4
Adults/subadults	Females	2(8)	2(3)	6(12)	2(2)
	Males	6(8)	5(5)	7(7)	7(7)
Juveniles	Females	4(5)	0(0)	0(2)	0(0)
	Males	1(2)	2(2)	0(0)	0(1)
Infants	Females/Males	0(2)	0(1)	0(3)	0(1)
	Total number of chimpanzees	13(25)	9(11)	13(24)	9(11)

Table B.2. Context duration (minutes) per individual per time periods. Sum, proportion, mean and SD are calculated across all individuals for each context. *Other corresponds to contexts that rarely occurred in the data, such as object manipulation or copulation.

	2013									2017								
ID	Feeding	Resting	Solitude	Grooming	Play	Vigilance	Locomotion	Aggression	Other	Feeding	Resting	Solitude	Grooming	Play	Vigilance	Locomotion	Aggression	Other
1	32.63	22.53	10.33	9.32	0.93	0.23	2.73	0.00	0.00	43.05	35.97	5.20	9.90	0.37	0.00	2.30	0.00	3.22
3	8.00	5.92	3.72	1.92	0.00	0.00	0.28	0.00	0.00	47.68	13.08	18.27	10.77	0.00	3.47	3.67	0.12	0.00
4	20.00	36.85	8.85	4.23	4.73	0.38	3.07	0.00	0.00	24.38	35.82	20.80	8.38	0.47	5.28	1.88	0.47	2.52
7	23.27	31.37	10.95	2.15	6.28	0.18	0.32	0.10	0.00	42.30	47.40	3.92	3.10	0.00	0.00	1.43	0.70	0.00
8	10.70	4.00	4.00	0.00	0.00	0.00	0.00	0.00	0.00	40.42	33.22	12.48	6.97	0.00	4.62	1.68	0.38	0.00
10	4.00	11.78	0.00	2.73	0.00	0.00	1.07	0.00	0.00	49.75	16.42	10.60	7.83	0.00	11.15	4.00	0.25	0.00
13	7.97	8.37	0.00	5.07	2.22	0.00	0.00	0.00	0.00	50.68	15.07	10.23	10.00	0.35	1.88	1.45	0.00	0.00
15	9.00	1.35	4.07	0.00	0.00	0.00	0.68	0.00	0.90	53.23	18.37	5.90	6.17	0.00	11.17	0.00	0.78	0.00
16	4.00	15.55	0.00	0.00	0.00	0.03	0.37	0.05	1.28	61.48	19.40	2.77	10.30	0.00	4.45	0.88	0.27	0.00
24	3.37	0.00	2.03	0.20	0.00	0.43	0.00	0.00	0.00	53.95	20.68	3.90	0.00	1.28	5.22	0.00	0.20	0.00
29	21.90	15.48	28.38	5.57	2.07	0.00	0.32	0.00	0.00	27.15	23.62	33.18	0.58	0.55	0.00	13.70	0.22	0.22
32	16.00	42.12	2.03	7.28	4.28	1.62	1.28	0.05	0.00	45.00	28.42	10.00	11.68	3.95	0.77	0.18	0.00	0.00
36	4.00	8.00	1.12	0.00	0.00	2.88	0.00	0.00	0.00	37.63	26.38	8.12	3.83	0.33	9.50	5.10	0.72	0.00
42	7.42	10.78	0.00	0.00	0.92	0.00	0.00	0.15	0.00	33.08	6.60	35.00	0.00	17.37	1.58	1.07	1.05	0.00
47	27.12	26.72	8.87	3.27	10.00	2.68	0.00	0.00	0.00	50.62	3.00	17.40	11.43	5.00	0.13	9.48	1.02	1.67
50	15.32	27.10	6.15	13.27	6.42	0.42	3.97	1.13	0.00	42.73	37.82	13.47	8.97	0.15	0.33	1.53	0.00	0.00
53	14.30	7.73	4.00	0.00	0.00	0.00	1.02	0.00	0.00	63.60	11.93	12.48	2.98	0.00	5.08	0.83	0.67	0.00
56	3.63	15.13	0.00	2.15	2.27	0.82	2.47	0.00	0.93	47.67	22.18	13.77	0.00	0.00	12.10	1.92	0.77	0.25
57	0.00	7.83	3.00	0.68	3.32	0.00	0.17	0.00	0.00	55.00	20.30	12.37	4.05	0.62	0.00	2.52	0.00	0.42
60	6.78	21.17	28.10	8.35	3.75	5.12	2.43	0.53	0.00	29.95	35.52	14.18	9.50	1.33	0.00	1.75	1.25	1.28
65	8.00	0.82	0.00	2.85	11.82	0.00	0.00	0.00	0.00	56.95	17.90	3.02	14.00	0.30	2.22	4.98	1.88	0.00
66	36.67	21.25	7.85	5.02	7.15	0.00	2.55	0.00	0.00	36.73	25.75	25.57	7.38	1.65	0.00	1.33	0.13	1.12

Grand Total	284.07	341.85	133.45	74.05	66.15	14.80	22.72	2.02	3.12	993.05	514.83	292.62	147.83	33.72	78.95	61.70	10.87	10.68
%	30.15	36.28	14.16	7.86	7.02	1.57	2.41	0.21	0.33	46.31	24.01	13.65	6.89	1.57	3.68	2.88	0.51	0.50
Mean	12.91	15.54	6.07	3.37	3.01	0.67	1.03	0.09	0.14	45.14	23.40	13.30	6.72	1.53	3.59	2.80	0.49	0.49
SD	10.13	11.82	8.02	3.62	3.48	1.30	1.24	0.26	0.37	10.65	10.96	8.96	4.27	3.77	4.07	3.24	0.49	0.91

Table B.3. Overview of the z-score per subject per behaviour within each context across the two time periods.

Context	ID	Body contact	Food begging	Food sharing	Groom	Playful contact	Chase	Display	Risky approach	Risky action	Rough action	Throw	Gaze	Gaze Approach	Gaze Manipulate	Gaze Touch	Escape	Pilo-erection	Rock	Self-scratch	Self-touch	Vigilant	Yawn
Feeding	1	0.45	-0.42	-0.34	0.22	-0.32	-0.34		-0.65	-0.68		-0.62	-0.26	-0.45	-0.44	-0.29				0.42	-0.14		-0.52
	3	-0.48	-0.42	-0.34	-0.33	-0.32	0.09		-0.48	0.95		0.03	-0.26	1.29	0.22	-0.29				-0.47	-0.75		-0.13
	4	1.19	-0.42	-0.34	1.87	-0.32	-0.34		0.71	-0.36		-0.62	-0.26	0.19	0.22	-0.29				-1.06	-0.83		-0.10
	7	0.24	-0.42	0.18	0.22	-0.32	-0.34		0.28	0.86		-0.62	-0.26	-0.45	-0.44	-0.29				0.31	0.03		-0.34
	8	-0.70	-0.42	-0.34	-0.33	-0.32	-0.34		-0.04	-0.68		-0.62	-0.26	-0.45	-0.44	-0.29				-0.82	-0.51		-0.52
	10	-0.25	-0.42	-0.34	-0.33	0.48	-0.34		-0.49	0.62		1.14	0.34	-0.23	-0.44	-0.29				-0.49	1.19		-0.52
	13	-0.70	-0.42	1.22	-0.33	-0.32	-0.34		-0.35	-0.36		1.25	-0.26	-0.45	-0.44	-0.29				0.55	-0.54		-0.33
	15	0.41	0.76	-0.34	-0.33	-0.32	0.07		0.12	-0.03		0.69	-0.26	-0.45	0.53	-0.29				-0.39	0.25		-0.15
	16	0.32	-0.42	0.21	0.22	-0.32	-0.34		-0.34	-0.03		1.93	-0.26	0.05	0.42	-0.29				-0.37	-0.47		-0.52
	24	-0.14	2.65	-0.34	-0.33	-0.32	-0.34		-0.01	-0.04		-0.19	2.00	-0.45	0.22	1.83				0.05	-0.67		0.96
	29	-0.70	-0.42	-0.34	-0.33	-0.32	-0.34		-0.65	-0.68		-0.19	-0.26	2.35	1.53	-0.29				0.56	2.20		-0.52
	32	0.08	-0.42	1.21	-0.33	0.89	-0.34		-0.35	-0.36		-0.62	-0.26	0.05	-0.24	-0.29				1.63	0.14		-0.52
	36	-0.59	1.11	-0.34	-0.33	-0.32	-0.34		1.10	-0.04		-0.62	-0.26	-0.45	-0.44	-0.29				-0.60	-0.04		-0.34
	42	-0.70	-0.42	0.69	-0.33	-0.32	-0.34		-0.65	-0.68		-0.39	-0.26	-0.45	-0.44	0.09				-0.54	-0.25		0.64

	47	0.54	-0.42	-0.34	-0.33	-0.32	0.65	-0.27	0.83	-0.62	-0.26	-0.45	-0.44	-0.29	0.23	0.15	-0.52			
	50	1.59	2.10	-0.34	-0.33	1.72	3.26	1.03	2.53	-0.62	-0.26	0.41	0.22	2.06	1.33	0.48	-0.15			
	53	-0.29	0.38	-0.34	0.26	0.29	-0.34	0.16	0.60	1.04	2.09	1.40	1.92	0.10	1.05	0.79	0.42			
	56	0.06	-0.42	-0.34	-0.33	-0.32	-0.34	0.45	-0.68	0.25	-0.26	-0.45	-0.44	-0.29	-0.94	-0.53	0.05			
	57	-0.04	-0.42	-0.34	-0.33	-0.32	-0.34	-0.50	-0.68	0.25	-0.26	-0.45	-0.44	-0.29	-0.82	-0.74	-0.34			
	60	-0.59	-0.42	-0.34	-0.33	-0.32	-0.34	-0.19	-0.36	-0.62	-0.26	-0.45	-0.44	-0.29	-0.54	-0.33	0.73			
	65	-0.24	-0.42	-0.34	0.22	1.94	1.59	-0.80	-0.04	0.04	0.31	0.41	0.22	0.82	0.22	-0.33	-0.52			
	66	0.46	-0.42	2.01	2.02	-0.32	-0.34	1.94	-0.68	-0.32	-0.26	-0.45	-0.44	-0.29	0.70	0.92	3.24			
Affiliative	1	-0.10			0.20	-0.44	-0.21		-0.23	-0.01	-0.38	-0.16	0.05	-0.25	-0.21	-0.92	-0.72	0.24	1.22	
	3	1.32			0.17	0.08	-0.21		-0.24	-0.80	-0.38	-0.37	-0.43	-0.25	-0.21	-0.68	-0.62	-0.26	-0.50	
	4	-0.44			2.11	1.40	-0.21		0.32	0.46	0.24	0.00	-0.02	-0.25	-0.21	0.64	-0.02	-0.27	-0.50	
	7	-0.55			-0.78	-0.44	-0.21		-0.44	1.33	1.43	-0.37	-0.43	-0.25	-0.21	-0.78	-0.87	-0.46	0.32	
	8	-0.49			-0.65	-0.44	-0.21		-0.44	-0.35	-0.38	-0.37	-0.43	-0.25	-0.21	-0.74	-0.87	-0.46	-0.24	
	10	-0.55			0.46	-0.14	-0.21		0.65	-0.60	-0.38	-0.37	0.86	-0.25	-0.21	-0.66	-0.72	-0.11	-0.50	
	13	0.62			-0.37	-0.16	-0.21		-0.25	-0.16	-0.38	-0.37	-0.43	-0.25	-0.21	0.52	-0.62	-0.46	-0.50	
	15	-0.55			-0.50	-0.44	-0.21		-0.44	0.07	-0.38	-0.37	-0.43	-0.25	-0.21	-0.28	-0.18	-0.46	-0.50	
	16	-0.42			-0.56	0.93	-0.21		0.70	0.93	1.00	1.52	-0.43	2.10	-0.21	0.57	0.67	0.65	-0.50	

	24	-0.55	0.63	-0.44	-0.21	-0.44	-0.03	-0.38	-0.37	-0.43	-0.25	-0.21	-0.65	-0.61	-0.46	-0.50
	29	1.31	-0.20	3.23	4.48	1.49	-0.11	2.10	0.93	3.55	1.99	-0.21	-0.38	0.06	-0.46	-0.11
	32	-0.55	0.08	-0.44	-0.21	-0.44	-0.32	-0.38	-0.37	-0.43	-0.25	-0.21	-0.11	-0.53	-0.46	-0.50
	36	-0.55	-0.87	-0.44	-0.21	-0.44	-0.25	-0.38	-0.37	-0.43	-0.25	-0.21	-0.92	-0.74	-0.46	-0.50
	42	-0.55	-0.95	-0.35	-0.21	-0.10	-0.80	-0.38	-0.37	-0.43	-0.25	2.14	-0.79	-0.22	-0.14	-0.50
	47	-0.55	0.88	0.32	-0.21	-0.13	0.07	0.03	-0.13	-0.27	-0.25	-0.21	0.07	-0.62	-0.46	0.25
	50	0.74	-0.05	-0.17	-0.21	1.91	2.53	0.84	0.29	-0.11	0.60	-0.21	0.46	0.40	-0.46	-0.50
	53	0.19	-0.14	-0.44	-0.21	-0.44	-0.73	-0.38	-0.30	0.02	-0.25	-0.21	-0.38	0.44	2.69	1.10
	56	-0.19	-0.11	-0.44	-0.21	-0.44	-0.45	-0.38	1.97	0.68	-0.25	-0.21	0.45	1.78	-0.46	0.00
	57	0.58	-0.21	-0.38	-0.21	-0.26	-0.24	-0.38	-0.37	-0.43	-0.25	-0.21	1.90	0.92	0.32	0.77
	60	1.36	0.46	-0.18	-0.21	0.22	0.16	0.06	0.45	0.60	-0.25	2.14	1.36	1.36	0.71	1.95
	65	-0.55	-0.39	-0.37	-0.21	-0.44	-0.62	-0.38	-0.10	-0.43	-0.25	-0.21	0.23	0.52	-0.27	-0.25
	66	0.39	0.80	-0.29	-0.21	-0.12	-0.08	-0.38	-0.37	-0.25	-0.25	-0.21	1.08	1.14	1.52	0.52
Resting	1	0.10	0.60	-0.45		-0.40	-0.36	-0.29	-0.35	-0.29			-0.61	-0.81	2.30	0.54
	3	0.02	-0.05	-0.45		-0.40	0.71	-0.29	1.55	1.06			0.15	0.33	-0.41	-0.10
	4	0.77	1.61	-0.13		-0.10	-0.23	-0.29	-0.42	-0.37			0.65	0.57	-0.06	0.73
	7	0.56	0.43	0.81		0.67	0.32	1.62	-0.04	-0.18			0.13	0.76	-0.06	0.87

8	-0.50	-0.56	-0.45	-0.40	-0.51	-0.29	-0.42	-0.37	-0.52	0.07	-0.33	-0.65
10	-0.42	1.44	-0.45	-0.40	-0.56	-0.29	-0.42	-0.37	-0.27	-0.58	0.21	0.19
13	-0.50	0.27	-0.45	-0.40	0.36	-0.29	-0.42	-0.27	-0.16	0.97	-0.33	0.18
15	-0.37	-0.56	-0.45	-0.40	0.80	-0.29	-0.42	-0.37	-0.01	-0.14	-0.09	-0.65
16	-0.42	-0.18	0.65	2.11	-0.51	-0.29	-0.42	-0.37	-0.13	-0.47	-0.08	-0.65
24	-0.50	-0.56	-0.45	-0.40	0.96	-0.29	-0.10	-0.05	-0.99	0.08	-0.41	1.01
29	-0.06	-0.56	-0.45	-0.08	2.24	2.06	3.02	3.88	0.98	1.26	-0.25	-0.08
32	1.38	-0.02	-0.45	-0.40	-0.23	-0.29	-0.17	-0.20	1.11	0.50	-0.33	-0.65
36	-0.50	-0.56	1.90	-0.40	-0.22	0.95	-0.02	-0.16	0.51	0.22	-0.22	-0.65
42	-0.50	-0.56	0.64	0.39	-0.67	-0.29	-0.42	-0.37	-0.83	-0.25	-0.41	-0.65
47	-0.50	0.31	-0.45	-0.40	-0.01	-0.29	-0.22	-0.37	1.20	-0.58	-0.14	0.87
50	1.23	-0.56	0.30	1.91	-0.67	0.63	1.46	0.48	0.65	0.08	-0.26	-0.65
53	-0.50	-0.56	0.32	0.21	-0.25	-0.29	-0.33	-0.31	-0.87	-0.29	-0.41	-0.65
56	-0.50	0.65	0.43	-0.40	-0.34	-0.29	-0.42	-0.37	-0.76	-0.71	-0.22	0.26
57	1.66	-0.56	-0.45	0.07	0.04	-0.29	-0.42	-0.37	-0.90	-0.88	-0.18	-0.65
60	0.22	0.77	0.93	0.02	-0.38	-0.29	-0.42	-0.08	0.76	-0.07	0.10	1.75
65	-0.50	-0.56	-0.45	-0.40	-0.51	-0.29	-0.35	-0.27	-1.07	-0.35	-0.41	-0.65

	66	-0.16		-0.26	-0.45		-0.40	0.01	-0.29	-0.30	-0.29		0.99	0.31	1.96	0.28
Solitude	1								-0.21	-0.37	-0.39	2.09	-0.20	-0.48	-0.24	-0.38
	3								-0.21	-0.16	0.11	-0.26	0.33	-0.49	-0.41	1.12
	4								-0.21	-0.37	-0.39	-0.26	-0.11	-0.49	0.99	-0.38
	7								-0.21	-0.37	-0.39	-0.26	-0.29	-0.48	0.57	-0.38
	8								-0.21	-0.37	-0.39	-0.26	-0.62	-0.69	-0.41	0.12
	10								-0.21	-0.37	-0.39	-0.26	0.09	0.28	-0.45	-0.38
	13								2.14	-0.15	-0.28	-0.26	0.12	0.58	-0.51	-0.38
	15								-0.21	-0.37	-0.39	-0.26	0.04	-0.47	2.29	-0.38
	16								-0.21	-0.37	-0.39	-0.26	-0.69	-0.74	-0.51	-0.38
	24								-0.21	-0.37	-0.39	-0.26	-0.69	-0.56	-0.51	-0.38
	29								-0.21	3.84	2.32	-0.26	2.38	2.07	0.09	0.42
	32								-0.21	-0.16	-0.17	-0.26	-0.69	-0.36	-0.51	-0.38
	36								-0.21	0.73	1.93	-0.26	-0.17	0.07	-0.42	-0.38
	42								-0.21	-0.37	-0.39	-0.26	-0.46	0.56	-0.51	-0.38
	47								-0.21	-0.37	-0.28	1.74	-0.33	-0.69	-0.51	0.12
	50								-0.21	-0.37	0.81	-0.26	1.02	2.13	-0.07	-0.38

53	-0.21	0.47	0.04	-0.26	-0.11	-0.26	-0.51	-0.38
56	-0.21	-0.13	-0.16	-0.26	-0.53	-0.52	-0.51	0.12
57	-0.21	-0.37	-0.39	1.08	-0.69	0.24	0.01	-0.38
60	2.14	0.71	0.34	-0.26	0.66	0.88	0.12	3.71
65	-0.21	-0.37	-0.39	-0.26	-0.69	-0.74	-0.51	-0.38
66	-0.21	-0.37	-0.39	-0.26	1.62	0.13	2.47	0.12

Table B.4. Additional GT analysis when considering the feeding (F) and resting (R) contexts. Estimated variance components for the G-study subject by context by time design and subject by time design as well as the G-coefficients per behaviour.

Behaviour	Body contact		Groom		Playful contact		Rough action		Gaze		Gaze Approach		Gaze Manipulate		Gaze Touch		Self-scratch		Self-touch		Vigilant	
Effect	Var	%	Var	%	Var	%	Var	%	Var	%	Var	%	Var	%	Var	%	Var	%	Var	%	Var	%
Subject	-0.018	-1.90	0.237	24.60	-0.137	-14.50	0.510	52.20	-0.130	-13.80	-0.084	-8.80	0.352	36.30	0.296	30.60	-0.003	-0.30	-0.071	-7.40	0.350	36.10
Context	0.006	0.60	0.015	1.60	0.000	-0.10	0.017	1.70	-0.008	-0.80	0.000	0.00	0.001	0.10	0.000	0.00	-0.007	-0.70	0.002	0.20	0.006	0.60
Time	-0.019	-2.00	0.010	1.00	-0.008	-0.80	0.022	2.20	-0.012	-1.30	0.000	0.00	-0.015	-1.50	-0.002	-0.20	-0.016	-1.70	-0.017	-1.70	0.008	0.90
Subject*Context	-0.130	-13.60	-0.330	-34.10	0.011	1.20	-0.370	-37.80	0.172	18.10	0.008	0.80	-0.016	-1.60	-0.004	-0.40	0.144	15.10	-0.042	-4.40	-0.123	-12.60
Subject*Time	0.427	44.70	-0.211	-21.90	0.170	17.90	-0.478	-48.90	0.272	28.70	0.008	0.80	0.327	33.70	0.049	5.10	0.359	37.70	0.366	38.50	-0.184	-18.90
Context*Time	-0.033	-3.40	-0.059	-6.10	-0.044	-4.60	-0.061	-6.20	-0.031	-3.30	-0.049	-5.10	-0.015	-1.60	-0.030	-3.10	-0.023	-2.40	-0.034	-3.60	-0.044	-4.50
Subject*Context*Time	0.722	75.70	1.303	135.00	0.957	100.90	1.337	136.80	0.685	72.30	1.068	112.30	0.336	34.70	0.657	68.00	0.500	52.40	0.747	78.50	0.957	98.50
G-coefficient	-0.06		0.81		-0.72		1.21		-0.50		-0.44		0.60		0.61		-0.01		-0.25		0.80	

Table B.5. Additional GT analysis when considering feeding (F), affiliative (GP), resting (R) and solitude (S) contexts. Estimated variance components for the G-study subject by context by time design and subject by time design as well as the G-coefficients per behaviour.

Context	F-GP-R		F-GP-R		F-GP-R		F-GP		F-GP-R		F-GP-R		F-GP-R-S	
Behaviour	Body contact		Groom		Playful contact		Chase		Rough action		Gaze		Gaze Approach	
Effect	Var	%	Var	%	Var	%	Var	%	Var	%	Var	%	Var	%
Subject	0.028	3.00	0.245	25.40	-0.098	-10.30	-0.074	-7.80	0.289	29.90	-0.137	-14.50	-0.074	-7.80
Context	0.006	0.70	0.014	1.50	-0.008	-0.80	-0.037	-3.90	0.011	1.20	-0.010	-1.10	-0.001	-0.10
Time	-0.005	-0.50	0.002	0.20	-0.002	-0.20	0.000	0.00	0.003	0.30	-0.009	-0.90	-0.007	-0.80
Subject*Context	-0.142	-14.80	-0.311	-32.20	0.177	18.60	0.810	85.10	-0.245	-25.40	0.226	23.90	0.021	2.20
Subject*Time	0.107	11.20	-0.034	-3.50	0.045	4.80	0.000	0.00	-0.072	-7.50	0.189	20.00	0.160	16.80
Context*Time	-0.046	-4.80	-0.050	-5.20	-0.040	-4.20	-0.012	-1.30	-0.047	-4.80	-0.033	-3.50	-0.041	-4.30
Subject*Context*Time	1.008	105.30	1.099	113.80	0.876	92.10	0.264	27.80	1.027	106.30	0.720	76.00	0.894	94.00
G-coefficients	0.14		0.80		-0.75		-0.19		0.84		-0.90		-0.60	

Context	F-GP-R-S		F-GP-R-S		F-GP		F-GP-R-S		F-GP-R-S		F-GP-R-S		GP-R-S	
Behaviour	Gaze Manipulate		Gaze Touch		Escape		Self-scratch		Self-touch		Vigilant		Yawning	
Effect	Var	%	Var	%	Var	%	Var	%	Var	%	Var	%	Var	%
Subject	0.280	29.00	0.404	41.60	-0.077	-8.10	0.159	16.60	-0.006	-0.60	0.304	31.30	0.340	35.10
Context	-0.001	-0.10	0.004	0.40	0.000	0.00	-0.002	-0.20	-0.005	-0.50	0.000	0.00	0.005	0.50
Time	-0.004	-0.50	0.001	0.10	-0.009	-0.90	0.002	0.20	-0.010	-1.00	0.004	0.40	0.003	0.30
Subject*Context	0.031	3.20	-0.083	-8.60	0.005	0.50	0.045	4.70	0.102	10.70	-0.003	-0.30	-0.114	-11.80
Subject*Time	0.096	10.00	-0.026	-2.70	0.193	20.30	-0.039	-4.10	0.211	22.10	-0.084	-8.70	-0.063	-6.50
Context*Time	-0.027	-2.80	-0.032	-3.30	-0.040	-4.20	-0.038	-3.90	-0.031	-3.30	-0.036	-3.70	-0.038	-3.90
Subject*Context*Time	0.592	61.30	0.704	72.40	0.878	92.40	0.834	86.80	0.693	72.60	0.785	80.90	0.837	86.30
G-coefficients	0.68		0.88		-0.32		0.62		-0.03		0.85		0.83	

Context	GP		S		F		F		F	
Behaviour	Pilo-erection		Rock		Food begging		Food sharing		Risky action	
Effect	Var	%	Var	%	Var	%	Var	%	Var	%
Subject	-0.048	-5.00	-0.068	-7.10	0.545	55.40	-0.106	-11.10	-0.055	-5.80
Time	-0.048	-5.00	-0.049	-5.10	-0.021	-2.10	-0.050	-5.30	-0.048	-5.00
Subject*Time	1.047	110.00	1.067	112.20	0.460	46.70	1.104	116.40	1.054	110.80
G-coefficients	-0.10		-0.15		0.70		-0.24		-0.12	

Table B.6. Sex analysis. Overview of the statistical analyses comparing males and females' personality scores using the Mann-Whitney U test. N=22.

Personality trait	Mann-Whitney U	Z	Exact Sig. (2-tailed)
Explorativeness	45.00	-0.76	0.470
Anxiety	45.00	-0.76	0.482
Boldness	37.00	-1.30	0.202
Sociability	43.00	-0.91	0.383

Table B.7. Age analysis. Overview of the statistical analyses correlating the subjects' age (Median age 2013-2017) and their personality trait scores using Spearman rank correlation. N=22.

Personality trait	r_s	Sig. (2-tailed)
Explorativeness	-0.35	0.115
Anxiety	-0.07	0.750
Boldness	-0.43*	0.048
Sociability	0.317	0.151

r_s: Spearman's rho; *p < 0.05.

C. Appendix 3. Supplementary material for Chapter 4

Table C.1. Subject representation and colony composition. Overview of the number of subjects per age and sex group per colony including for the naturalistic and rating approach, respectively. The number of subjects that were included in the experiment approach (snake test and novel object test, respectively) is shown in square parentheses. The number of chimpanzees that made up each colony is shown in parenthesis. Adults/subadults were over 8 years old, juveniles were 4-8 years old, and infants were less than 4 years old.

Age group	Sex	Colony 1	Colony 4
Adult/subadult	Female	6; 7[2,5] (12)	2; 2[1,1] (2)
	Male	7; 7[2,3] (7)	7; 7[5,4] (7)
Juvenile	Female	0; 0[0,0] (2)	0; 0[0,0] (0)
	Male	0; 0[0,0] (0)	1; 1[0,0] (1)
Infant	Female/Male	0; 0[0,0] (3)	0; 0[0,0] (1)
	Total number of chimpanzees	13; 14[4,8] (24)	10; 10[6,5] (11)

Table C.2. Questionnaire combining the adjective and behaviour-descriptive verb items. Each keeper received one hard copy to rate the personality of each chimpanzee using the 5-point Likert scale (almost never, rarely, sometimes, often, very often).

Rater's name:

Instructions: There are no right or wrong answers. Please answer what you think is the most appropriate for this chimp in this week, read every statement carefully and think how well it applies to the particular chimpanzee in comparison with others.

For how long have you known this chimpanzee?

Items	**Almost never**	**Rarely**	**Sometimes**	**Often**	**Very often**
_____ is curious.					
_____ is eager to learn.					
_____ is bold.					
_____ is social.					
When waiting for food, _____ scratches him/herself or walks back-and-forth.					
_____ is cheeky.					
_____ is friendly to other chimps.					
When others fight, _____screams or runs away.					
When playing with stronger chimps, _____ plays rough and chases them.					
If there is a flashing toy in the enclosure, _____approaches and touches it.					
When relaxing, _____ is nearby other chimps.					
_____ is easily alarmed.					
If there is a new object (like when the water or juice fountain was built) in the enclosure, _____ touches it.					
When eating next to stronger chimps, _____ takes food from them.					
During feeding,_____ sits nearby other chimps.					
If there is a music box in the enclosure, _____ approaches and touches it.					
_____ is anxious.					

Table C.3. Rate of behaviour per individual per experimental session for both the snake and novel object test.
"T1/2": Experimental test 1 and test 2.

ID	Snake test										Novel object test							
	Approach		Move away		Straight		Gaze		Stress-related		Leave		Straight		Gaze		Stress-related	
	T1	T2	T1	T2	T1	T2	T1	T2	T1	T2	T1	T2	T1	T2	T1	T2	T1	T2
7	0.00	0.00	0.00	0.00	1.09	2.18	3.82	3.82	1.64	2.18	0.00	0.00	0.00	0.00	2.00	5.00	0.00	0.50
8	0.00	0.00	0.55	0.55	2.18	8.73	3.82	3.82	4.91	3.82	0.00	0.00	10.00	10.75	2.50	4.78	0.50	0.00
9	1.09	0.00	0.55	0.00	3.82	7.09	3.82	5.45	7.64	0.55	0.00	0.00	8.00	5.00	7.00	3.50	1.50	0.50
10	-	-	-	-	-	-	-	-	-	-	0.55	0.50	5.46	1.50	3.27	3.00	0.00	0.00
24	-	-	-	-	-	-	-	-	-	-	0.00	0.00	2.50	4.00	3.00	5.50	3.50	6.00
29	1.09	0.00	0.00	0.00	2.73	2.73	6.55	2.18	0.55	0.55	0.50	0.50	2.00	3.00	4.00	5.50	2.00	0.50
32	0.00	0.00	0.00	0.00	4.91	1.64	6.00	4.91	0.55	0.00	-	-	-	-	-	-	-	-
36	0.00	1.09	1.64	0.00	6.00	1.09	8.18	7.64	5.45	2.18	0.00	0.00	1.50	6.26	5.00	2.41	0.00	2.41
42	-	-	-	-	-	-	-	-	-	-	0.00	0.00	1.50	4.50	3.50	5.50	0.00	2.00
47	1.09	0.00	0.00	0.00	4.91	2.73	4.91	4.36	0.00	0.00	0.00	0.00	6.50	7.50	3.50	4.50	0.00	2.50
53	0.70	1.09	0.70	0.55	7.71	1.64	4.21	8.73	2.10	0.55	0.00	0.00	9.50	1.50	3.50	8.00	1.50	0.50
57	-	-	-	-	-	-	-	-	-	-	0.00	0.00	7.00	2.00	9.00	7.50	5.00	5.00
60	0.00	0.00	0.55	0.00	2.73	4.36	8.73	9.27	2.73	1.09	0.00	0.00	1.00	2.00	1.00	3.00	0.00	0.50
66	0.55	0.55	0.55	0.00	3.27	2.18	4.91	8.18	0.00	0.00	0.00	0.00	4.00	2.50	3.50	4.00	0.00	0.00

Table C.4. Additional GT analysis. Estimated variance components for the G-study subject by time design as well as the G-coefficients per personality trait for both SN and NO experimental tests.

		Anxiety		Boldness		Exploravitiveness	
Experimental test	Effect	Var	%	Var	%	Var	%
SN	Subject	0.504	52.80	-0.558	-32.3	0.059	3.50
	Time	-0.050	-5.20	-0.254	-14.7	-0.181	-10.70
	Subject*Time	0.499	52.40	2.536	147	1.814	107.20
	G-coefficient	0.669		-0.79		0.06	
NO	Subject	0.702	71.90	-	-	2.589	72.00
	Time	-0.023	-2.30	-	-	-0.084	-2.30
	Subject*Time	0.297	30.50	-	-	1.090	30.30
	G-coefficient	0.83				0.83	

D. Appendix 4. Supplementary material for Chapter 5

Table D.1. Context duration (minutes) per individual per time periods. Sum, proportion, mean and SD are calculated across all individuals for each context. *Other included contexts that rarely occurred (e.g., object manipulation or copulation).

	2016								2017							
ID	Feeding	Solitude	Play	Resting	Locomotion	Aggression	Vigilance	Other	Feeding	Solitude	Play	Resting	Locomotion	Aggression	Vigilance	Other
1	42.25	43.00	4.32	0.57	1.08	0.00	0.00	1.18	26.22	40.67	2.68	3.47	3.23	1.70	0.00	0.28
2	16.00	18.67	15.98	1.27	0.93	0.00	0.00	2.63	21.78	42.12	5.43	2.60	4.88	0.00	0.00	1.93
3	41.97	22.07	6.20	7.12	0.92	0.00	0.00	0.62	16.53	15.50	16.87	7.17	3.37	0.27	0.00	0.82
4	9.93	14.05	3.17	3.77	0.68	0.00	0.00	0.00	15.98	16.45	15.78	5.85	1.02	0.43	0.00	2.93
5	28.47	39.08	11.23	5.35	1.53	0.17	0.00	0.50	18.75	14.63	9.42	5.00	0.58	0.30	0.00	0.00
6	49.05	27.73	1.03	1.73	1.85	0.00	0.00	2.65	26.90	33.03	9.80	10.48	2.68	0.27	0.00	1.55
7	16.12	16.68	0.00	0.05	0.00	0.00	0.00	0.00	11.48	10.10	0.15	0.75	0.00	0.00	0.00	0.00
8	38.18	13.28	7.08	0.52	0.87	0.00	0.00	0.38	31.32	23.72	17.92	3.48	0.82	0.00	0.62	0.23
9	22.87	29.83	6.97	6.75	3.62	0.00	0.00	4.50	10.38	31.17	5.28	2.53	0.32	0.00	0.00	0.00
10	31.13	18.60	4.23	4.68	1.98	0.00	0.00	1.83	37.52	28.98	5.13	2.27	1.03	0.45	0.00	0.80
11	49.17	27.25	1.77	6.22	1.58	0.00	0.00	3.03	30.18	25.92	11.07	4.03	1.33	0.00	0.57	0.85
12	36.23	10.53	21.03	3.65	2.68	0.00	0.00	0.88	27.48	30.33	10.60	2.52	3.73	0.00	0.00	1.12
13	48.97	21.73	5.23	7.15	1.10	0.00	0.00	5.83	23.78	46.67	0.20	6.48	0.93	0.00	0.00	0.00
14	19.43	17.42	0.38	1.33	0.18	0.00	0.00	0.00	21.52	19.53	0.00	0.38	1.13	0.00	0.00	0.00
15	22.35	29.72	2.23	1.13	0.90	0.00	0.00	1.93	27.12	31.68	1.43	11.92	2.30	0.37	0.00	4.98
16	0.00	7.82	0.95	0.15	0.00	0.00	0.00	0.00	22.18	35.20	0.00	2.25	1.67	0.20	0.00	3.00
17	16.95	26.33	0.00	3.00	0.18	0.00	0.00	0.00	20.87	27.32	0.48	1.48	1.22	0.00	0.00	0.00
18	19.78	9.52	9.17	2.28	0.60	0.00	0.00	4.73	22.12	28.30	10.48	11.57	2.05	0.00	0.00	0.00

19	32.75	11.78	3.65	5.75	0.48	0.00	0.00	1.23	39.17	11.73	12.62	10.10	2.67	0.00	0.00	3.00
20	48.73	28.08	4.83	8.32	0.25	0.00	0.00	2.52	14.63	9.72	1.83	3.25	0.63	0.00	0.00	0.00
Sum	590.33	433.18	109.47	70.78	21.43	0.17	0.00	34.47	465.92	522.77	137.18	97.58	35.60	3.98	1.18	21.50
%	46.86	34.38	8.69	5.62	1.70	0.01	0.00	2.74	36.24	40.66	10.67	7.59	2.77	0.31	0.09	1.67
Mean	29.52	21.66	5.47	3.54	1.07	0.01	0.00	1.72	23.30	26.14	6.86	4.88	1.78	0.20	0.06	1.08
SD	14.57	9.66	5.46	2.70	0.93	0.04	0.00	1.75	7.73	10.83	6.07	3.61	1.29	0.39	0.18	1.41

Table D.2. Additional GT analysis when considering the feeding (F), play (P) and solitude (S) contexts. Estimated variance components for the G-study subject by context by time design and subject by time design as well as the G-coefficients per behavioural action. The analysis is based on 11 subjects.

Context	FP		FP		FS		FPS		FPS		FS		FP		FP		FP	
Behaviour	Body contact		Playful contact		Escape		Self-scratch		Self-touch		Vigilant		Chase		Rough action		Risky action	
Effect	Var	%	Var	%	Var	%	Var	%	Var	%	Var	%	Var	%	Var	%	Var	%
Subject	0.021	3.30	0.111	10.60	0.266	28.30	0.092	8.60	0.564	42.00	-0.131	-17.20	-0.146	-14.20	0.299	34.30	0.147	19.10
Context	-0.053	-8.50	-0.047	-4.50	0.110	11.70	0.016	1.50	0.049	3.70	-0.012	-1.60	0.051	4.90	0.006	0.70	-0.059	-7.60
Time	-0.042	-6.80	-0.032	-3.00	0.036	3.90	0.013	1.20	0.011	0.80	-0.025	-3.30	0.021	2.00	-0.045	-5.20	-0.075	-9.70
Subject*Con	0.079	12.70	0.416	39.80	-0.247	-26.40	0.004	0.30	-0.315	-23.40	0.078	10.20	-0.081	-7.80	-0.147	-16.80	-0.083	-10.80
Subject*Tim	0.162	26.10	0.312	29.90	-0.307	-32.80	-0.074	-6.90	-0.208	-15.40	0.071	9.30	0.301	29.20	0.261	29.90	0.093	12.00
Context*Tim	0.058	9.30	0.034	3.20	-0.108	-11.50	-0.097	-9.10	-0.098	-7.30	-0.037	-4.90	-0.086	-8.30	0.063	7.20	0.072	9.40
Subject*Con	0.397	63.90	0.250	24.00	1.189	126.70	1.114	104.30	1.340	99.70	0.819	107.40	0.969	94.20	0.434	49.90	0.676	87.70
G-coefficient	0.09		0.21		0.93		0.38		0.98		-0.89		-0.71		0.64		0.46	

Context	FP		FPS		FS		FPS		F		F		S		F	
Behaviour	Resist		Gaze		Gaze Manipulate		Gaze Touch		Food begging		Food sharing		Yawn		Gaze Approach	
Effect	Var	%	Var	%	Var	%	Var	%	Var	%	Var	%	Var	%	Var	%
Subject	-0.257	-25.70	0.157	13.40	0.009	0.80	0.206	15.30	0.504	60.30	0.102	38.90	-0.042	-4.30	0.027	3.20
Context	-0.013	-1.30	0.032	2.70	0.020	1.70	0.020	1.50								
Time	-0.020	-2.00	0.054	4.60	-0.044	-3.70	-0.037	-2.80	0.005	0.60	0.020	7.80	-0.058	-6.00	-0.001	-0.10
Subject*Con	0.201	20.10	-0.145	-12.30	-0.274	-23.30	-0.005	-0.40								
Subject*Tim	0.091	9.10	-0.071	-6.10	0.471	40.10	0.351	26.00	0.327	39.10	0.139	53.30	1.062	110.30	0.829	96.90
Context*Tim	-0.062	-6.20	-0.098	-8.40	-0.078	-6.70	-0.064	-4.70								
Subject*Con	1.120	112.00	1.247	106.00	1.071	91.10	0.877	65.10								
G-coefficient	-1.52		0.56		0.02		0.39		0.76		0.594		-0.09		0.061	

Table D.3. Additional GT analysis when considering feeding (F), play (P), Resting (R) and solitude (S) contexts. Estimated variance components for the G-study subject by context by time design and subject by time design as well as the G-coefficients per behavioural action. The analysis is based on 20 subjects.

Context	F-P-R		F-R		F-P		F-S		F-P-R-S		F-P-R-S		F-R-S		F-P		F-P-R	
Behaviour	Body contact		Food begging		Playful contact		Escape		Self-scratch		Self-touch		Vigilant		Chase		Rough action	
Effect	Var	%	Var	%	Var	%	Var	%	Var	%	Var	%	Var	%	Var	%	Var	%
Subject	0.143	15.00	0.057	6.20	0.169	18.20	0.058	6.30	0.161	16.70	0.307	31.80	-0.065	-6.90	-0.135	-14.80	0.233	24.20
Context	0.001	0.10	-0.002	-0.20	-0.011	-1.20	0.003	0.30	-0.003	-0.30	-0.001	-0.10	0.000	0.00	0.000	0.00	0.003	0.30
Time	0.009	0.90	-0.007	-0.80	-0.007	-0.80	0.007	0.80	-0.002	-0.20	0.004	0.40	-0.016	-1.70	-0.010	-1.10	-0.006	-0.60
Subject*Context	-0.028	-2.90	0.041	4.40	0.233	25.10	-0.065	-7.10	0.062	6.50	0.019	2.00	-0.007	-0.80	-0.003	-0.30	-0.055	-5.70
Subject*Time	-0.175	-18.30	0.162	17.60	0.158	17.00	-0.146	-15.80	0.040	4.10	-0.082	-8.50	0.317	33.50	0.209	23.00	0.116	12.10
Context*Time	-0.053	-5.50	-0.033	-3.60	-0.019	-2.10	-0.053	-5.70	-0.037	-3.90	-0.038	-3.90	-0.038	-4.00	-0.042	-4.60	-0.035	-3.70
Subject*Context*Time	1.059	110.80	0.705	76.50	0.405	43.70	1.117	121.30	0.738	77.00	0.757	78.40	0.756	79.80	0.893	97.90	0.706	73.40
G-coefficient	0.64		0.17		0.36		0.25		0.56		0.84		-0.30		-0.71		0.60	

Context	F-P-R		F-P		F-P-R-S		F-R-S		F-P-R-S		F		S		F	
Behaviour	Risky action		Resist		Gaze		Gaze Manipulate		Gaze Touch		Food sharing		Yawning		Gaze Approach	
Effect	Var	%	Var	%	Var	%	Var	%	Var	%	Var	%	Var	%	Var	%
Subject	0.158	16.50	-0.139	-15.20	0.073	7.70	0.080	8.40	0.175	18.20	-0.029	-3.20	-0.092	-10.20	-0.101	-11.20
Context	0.007	0.80	-0.003	-0.30	0.002	0.20	-0.003	-0.30	0.001	0.10						
Time	-0.002	-0.20	-0.002	-0.20	0.001	0.10	-0.009	-0.90	-0.005	-0.50	-0.047	-5.20	-0.050	-5.50	-0.050	-5.60
Subject*Context	-0.144	-15.10	0.055	6.00	-0.039	-4.10	0.062	6.50	-0.023	-2.40						
Subject*Time	0.036	3.70	0.046	5.10	-0.024	-2.50	0.174	18.20	0.096	10.10	0.980	108.30	1.042	115.70	1.052	116.80
Context*Time	-0.048	-5.00	-0.047	-5.20	-0.049	-5.20	-0.034	-3.60	-0.038	-3.90						
Subject*Context*Time	0.952	99.20	1.003	109.90	0.990	103.70	0.685	71.70	0.752	78.40						
G-coefficient	0.55		-0.86		0.42		0.27		0.56		-0.06		-0.22		-0.24	

Table D.4. Estimated variance components for the G-study subject by context by time design per behavioural action for each age category when considering the feeding (F), play (P), resting (R), solitude (S) contexts. The analysis is based on 20 subjects.

	Context	F-P-R		F-R		F-P		F-S		F-P-R-S		F-P-R-S		F-R-S	
Age category	Behaviour	Body contact		Food begging		Playful contact		Escape		Self-scratch		Self-touch		Vigilant	
	Effect	Var	%	Var	%	Var	%	Var	%	Var	%	Var	%	Var	%
Juvenile	Subject	0.23	9.60	0.06	6.70	0.06	4.70	0.33	15.20	0.32	20.90	0.28	27.70	0.00	0.00
	Context	0.08	3.50	0.19	20.30	0.00	0.00	0.07	2.90	0.00	0.00	0.00	0.00	0.00	0.00
	Time	0.05	2.30	0.00	0.00	0.00	0.00	0.16	7.00	0.02	1.10	0.02	2.10	0.00	0.00
	Subject*Context	0.00	0.00	0.00	0.00	0.28	23.00	0.00	0.00	0.00	0.00	0.04	3.80	0.09	7.50
	Subject*Time	0.00	0.00	0.17	18.70	0.31	25.40	0.00	0.00	0.00	0.00	0.00	0.00	0.25	21.40
	Context*Time	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.08	6.60
	Subject*Context*Time	2.00	84.70	0.51	54.30	0.57	47.00	1.65	74.80	1.19	78.10	0.68	66.50	0.76	64.50
Adolescent	Subject	0.04	5.50	0.00	0.00	0.00	0.00	0.00	0.00	0.15	17.20	0.08	6.10	0.00	0.00
	Context	0.09	12.80	0.06	5.30	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Time	0.00	0.00	0.07	6.00	0.00	0.00	0.13	7.30	0.00	0.00	0.02	1.30	0.00	0.00
	Subject*Context	0.04	5.70	0.24	19.90	0.51	42.40	0.03	1.60	0.08	10.00	0.25	18.50	0.00	0.00
	Subject*Time	0.01	1.40	0.00	0.00	0.18	15.10	0.00	0.00	0.18	21.00	0.00	0.00	0.94	47.40
	Context*Time	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	2.60	0.00	0.00	0.11	5.60
	Subject*Context*Time	0.54	74.50	0.82	68.80	0.52	42.50	1.58	91.10	0.42	49.30	0.99	74.00	0.93	47.10
Adult	Subject	0.00	0.00	0.83	32.20	0.00	0.00	0.00	0.00	0.00	0.00	0.78	46.90	0.00	0.00
	Context	0.00	0.00	0.14	5.50	0.02	31.70	0.00	0.00	0.00	0.00	0.05	3.00	0.00	0.00
	Time	0.00	0.00	0.21	8.40	0.00	0.00	0.00	0.00	0.00	0.00	0.03	1.90	0.00	0.00
	Subject*Context	0.25	37.80	0.00	0.00	0.02	22.20	0.36	67.60	0.31	36.40	0.00	0.00	0.07	39.60
	Subject*Time	0.16	23.60	0.19	7.50	0.00	0.00	0.00	0.00	0.29	34.00	0.00	0.00	0.00	0.00
	Context*Time	0.01	1.70	0.00	0.00	0.02	29.40	0.01	2.20	0.00	0.00	0.00	0.00	0.00	0.00
	Subject*Context*Time	0.25	37.00	1.19	46.50	0.01	16.70	0.16	30.20	0.25	29.60	0.80	48.20	0.11	60.40

Age category	Context	F-P		F-P-R		F-P		F-P-R		F-P-R-S		F-R-S		F-P-R-S	
	Behaviour	Chase		Rough action		Resist		Risky action		Gaze		Gaze Manipulate		Gaze Touch	
	Effect	Var	%	Var	%	Var	%	Var	%	Var	%	Var	%	Var	%
Juvenile	Subject	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.20	0.10	7.20	0.11	10.00	0.11	7.60
	Context	0.00	0.00	0.07	5.40	0.00	0.00	0.05	4.70	0.04	2.90	0.03	2.30	0.00	0.00
	Time	0.00	0.00	0.02	1.40	0.00	0.00	0.01	0.90	0.00	0.00	0.00	0.00	0.00	0.00
	Subject*Context	0.11	7.70	0.00	0.00	0.32	26.20	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Subject*Time	0.20	14.20	0.16	11.90	0.31	25.30	0.00	0.00	0.00	0.00	0.12	10.10	0.25	16.90
	Context*Time	0.00	0.00	0.00	0.00	0.06	4.90	0.00	0.00	0.00	0.00	0.02	1.30	0.00	0.00
	Subject*Context*Time	1.12	78.10	1.08	81.20	0.53	43.60	1.09	94.20	1.30	89.90	0.87	76.20	1.12	75.50
Adolescent	Subject	0.00	0.00	0.54	34.80	0.56	27.90	0.04	2.80	0.00	0.00	0.00	0.00	0.13	15.20
	Context	0.00	0.00	0.08	4.80	0.09	4.50	0.05	3.70	0.00	0.00	0.00	0.00	0.00	0.00
	Time	0.00	0.00	0.08	5.30	0.00	0.00	0.06	4.10	0.00	0.00	0.00	0.00	0.00	0.00
	Subject*Context	0.05	2.70	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.68	33.20	0.05	6.50
	Subject*Time	0.86	49.40	0.00	0.00	0.16	8.10	0.00	0.00	0.00	0.00	0.57	27.90	0.03	3.80
	Context*Time	0.17	9.80	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.11	5.40	0.00	0.00
	Subject*Context*Time	0.66	38.00	0.86	55.10	1.19	59.40	1.31	89.40	1.08	100.00	0.68	33.50	0.62	74.50
Adult	Subject	0.14	15.50	0.02	6.10	0.01	1.50	0.00	0.00	0.02	4.30	0.00	0.00	0.02	3.80
	Context	0.04	4.10	0.03	8.50	0.09	9.50	0.28	13.70	0.03	6.40	0.00	1.70	0.05	12.10
	Time	0.00	0.00	0.00	0.00	0.01	0.60	0.00	0.00	0.00	0.00	0.00	0.00	0.01	1.50
	Subject*Context	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.90	0.00	0.00	0.00	0.00
	Subject*Time	0.00	0.00	0.18	58.50	0.00	0.00	0.00	0.00	0.09	19.20	0.00	0.00	0.00	0.00
	Context*Time	0.00	0.00	0.00	0.10	0.00	0.00	0.00	0.00	0.02	4.00	0.01	20.20	0.00	0.00
	Subject*Context*Time	0.73	80.40	0.08	26.80	0.84	88.40	1.77	86.30	0.31	65.10	0.05	78.00	0.36	82.60

E. Ethics forms

to adhere to these, or respond to them, unless they

Science Faculty Ethics Committee

Science Faculty Office
University of Portsmouth
St Michael's Building
White Swan Road
PORTSMOUTH
PO1 2DT

023 9284 3379
ethics-sci@port.ac.uk

16 May 2017

Ms Hélène Chotard
Department of Psychology
University of Portsmouth

Helene.chotard@port.ac.uk

FAVOURABLE ETHICAL OPINION – WITH CONDITIONS

Study Title: Personality Traits Of Chimpanzees: Rating Methods To Validate A New Approach.

Reference Number: SFEC 2017-049

Date Submitted: 2 May 2017

Thank you for submitting your application to the Science Faculty Ethics Committee (SEFC) for ethical review in accordance with current procedures.

I am pleased to inform you that SFEC was content to grant a favourable ethical opinion of the above research on the basis described in the submitted documents listed at Annex A, and subject to standard general conditions (*See Annex B)*, and the following specific conditions:

Conditions

A. Section 12.3: Please add the retention of consent forms for 30 years, in line with the University of Portsmouth Retention Schedule.

B. The debrief needs to name the first supervisor, Head of Department, and University Complaints Officer in accordance with the University complaints procedure.

Advisory Note(s)[1]

1. The application would benefit from a thorough proofread. It was also suggested that the author might consider writing in the third person rather than the first person.

2. Section 12.2: The use of a more secure data retention system was suggested. Rather than a password protected hard drive, which is subject to malfunction and loss, alternatives could be cloud based system, or preferably the university N drive.

[1] to adhere to these, or respond to them, unless they

Please resubmit an updated application form incorporating the changes as per the above conditions for the final SFEC records on this application.

[1] These comments are given in good faith and it is hoped they are accepted as such. The PI does not need to adhere to these, or respond to them, unless they

If you would find it helpful to discuss any of the matters raised above or seek further clarification from a member of the Committee, you are welcome to contact ethics-sci@port.ac.uk who will circulate your queries to SFEC

Please note that the favourable opinion of SFEC does not grant permission or approval to undertake the research. Management permission or approval must be obtained from any host organisation, including the University of Portsmouth or supervisor, prior to the start of the study.

Wishing you every success in your research

Dr John Crossland
Vice Chair Science Faculty Ethics Committee

Annexes

A - Documents reviewed
B - After ethical review - Guidance for researchers

Information:

Dr Marina Davila-Ross - PhD Supervisor
Dr Jérôme Micheletta - PhD Supervisor
Prof Kim Bard - PhD Supervisor
Holly Shawyer - Faculty Administrator

Statement of compliance

SFEC is constituted in accordance with the Governance Arrangements set out by the University of Portsmouth

After Ethical Review

If unfamiliar, please consult the advice After Ethical Review (Annex B), which gives detailed guidance on reporting requirements for studies with a favourable opinion, including, notifying substantial amendments, notification of serious breaches of the protocol, progress reports and notifying SFEC of the end of the study.

Feedback

You are invited to give your view of the service that you have received from the Science Faculty Ethics Committee. If you wish to make your views known please contact the administrator at ethics-sci@port.ac.uk

ANNEX A Documents reviewed

The documents ethically reviewed for this application

Document	*Version*	*Date*
A-2017-049 CHOTARD PI submission email		2 May 2017
B-2017-049 CHOTARD Ethics - Helene Chotard - Version 2 - Amended version	V1	2 May 2017
C-2017-049 CHOTARD Peer-Reviewed version - Version 1 - Ethics - Helene Chotard		2 May 2017

ANNEX B - After ethical review - Guidance for researchers

1. This Annex sets out important guidance for researchers with a favourable opinion from a University of Portsmouth Ethics Committee. Please read the guidance carefully. A failure to follow the guidance could lead to the committee reviewing and possibly revoking its opinion on the research.

2. It is assumed that the research will commence within 1 year of the date of the favourable ethical opinion or the start date stated in the application, whichever is the latest.

3. The research must not commence until the researcher has obtained any necessary management permissions or approvals – this is particularly pertinent in cases of research hosted by external organisations. The appropriate head of department should be aware of a member of staff's research plans.

4. If it is proposed to extend the duration of the study beyond that stated in the application, the Ethics Committee must be informed.

5. Any proposed substantial amendments must be submitted to the Ethics Committee for review. A substantial amendment is any amendment to the terms of the application for ethical review, or to the protocol or other supporting documentation approved by the Committee that is likely to affect to a significant degree:

(a) the safety or physical or mental integrity of participants
(b) the scientific value of the study
(c) the conduct or management of the study.

5.1 A substantial amendment should not be implemented until a favourable ethical opinion has been given by the Committee.

6. Researchers are reminded of the University's commitments as stated in the Concordat to Support Research Integrity viz:

- maintaining the highest standards of rigour and integrity in all aspects of research
- ensuring that research is conducted according to appropriate ethical, legal and professional frameworks, obligations and standards
- supporting a research environment that is underpinned by a culture of integrity and based on good governance, best practice and support for the development of researchers
- using transparent, robust and fair processes to deal with allegations of research misconduct should they arise
- working together to strengthen the integrity of research and to reviewing progress regularly and openly

7. In ensuring that it meets these commitments the University has adopted the UKRIO Code of Practice for Research. Any breach of this code may be considered as misconduct and may be investigated following the University Procedure for the Investigation of Allegations of Misconduct in Research. Researchers are advised to use the UKRIO checklist as a simple guide to integrity.

31 October 2016

University of Portsmouth

Professor Matt Guille

School of Biological Sciences

King Henry Building

Dear Dr Chotard,

RE: Ethics submission – Personality traits in great apes.

Approval of project by the Animal Welfare and Ethical Review Body (AWERB)

I am very happy to confirm that we were able to fast track your application and that the AWERB gave its approval for your proposal concerning work within the above project.

The AWERB uses UK Home Office guidelines on the Animals (Scientific Procedures) Act 1986 when assessing proposals and adheres to the regulations of the European Directive 2010/63/EU. Your project does not require a Home Office Project Licence as the animals are not subjected to procedures that have potential for harm and suffering. We are confident that the attached proposal demonstrates appropriate consideration of the Three Rs and animal welfare. Please use this letter as confirmation of ethical approval from AWERB, University of Portsmouth. Please use the number 1216A as confirmation of the successful review.

Yours sincerely,

MJ Guille PhD FSB
Professor of Developmental Genetics and Chair, AWERB

F. Form UPR 16

FORM UPR16

Research Ethics Review Checklist

Please include this completed form as an appendix to your thesis (see the Research Degrees Operational Handbook for more information

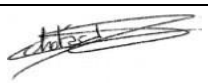
UNIVERSITY OF PORTSMOUTH

Postgraduate Research Student (PGRS) Information				**Student ID:**	726139
PGRS Name:	CHOTARD Hélène				
Department:	Psychology	**First Supervisor:**	Dr Marina Davila-Ross		
Start Date: (or progression date for Prof Doc students)	**01/10/2016**				
Study Mode and Route:	Part-time ☐	MPhil ☐	MD ☐		
	Full-time ☒	PhD ☒	Professional Doctorate ☐		
Title of Thesis:	Measuring Personality in Great Apes: Special Focus on Naturally Occurring Contexts				
Thesis Word Count: (excluding ancillary data)	44,258				

If you are unsure about any of the following, please contact the local representative on your Faculty Ethics Committee for advice. Please note that it is your responsibility to follow the University's Ethics Policy and any relevant University, academic or professional guidelines in the conduct of your study

Although the Ethics Committee may have given your study a favourable opinion, the final responsibility for the ethical conduct of this work lies with the researcher(s).

UKRIO Finished Research Checklist: (If you would like to know more about the checklist, please see your Faculty or Departmental Ethics Committee rep or see the online version of the full checklist at: http://www.ukrio.org/what-we-do/code-of-practice-for-research/)	
a) Have all of your research and findings been reported accurately, honestly and within a reasonable time frame?	YES ☒ NO ☐
b) Have all contributions to knowledge been acknowledged?	YES ☒ NO ☐
c) Have you complied with all agreements relating to intellectual property, publication and authorship?	YES ☒ NO ☐
d) Has your research data been retained in a secure and accessible form and will it remain so for the required duration?	YES ☒ NO ☐
e) Does your research comply with all legal, ethical, and contractual requirements?	YES ☒ NO ☐

Candidate Statement:	
I have considered the ethical dimensions of the above named research project, and have successfully obtained the necessary ethical approval(s)	
Ethical review number(s) from Faculty Ethics Committee (or from NRES/SCREC):	1216A; SFEC 2017-049
If you have *not* submitted your work for ethical review, and/or you have answered 'No' to one or more of questions a) to e), please explain below why this is so:	
Signed *(PGRS)*: 	**Date:** 30/10/2019